

# AN A POSTERIORI ERROR ANALYSIS FOR AN OPTIMAL CONTROL PROBLEM INVOLVING THE FRACTIONAL LAPLACIAN\*

HARBIR ANTIL<sup>†</sup> AND ENRIQUE OTÁROLA<sup>‡</sup>

**Abstract.** In a previous work, we introduced a discretization scheme for a constrained optimal control problem involving the fractional Laplacian. For such a control problem, we derived optimal a priori error estimates that demand the convexity of the domain and some compatibility conditions on the data. To relax such restrictions, in this paper, we introduce and analyze an efficient and, under certain assumptions, reliable a posteriori error estimator. We realize the fractional Laplacian as the Dirichlet-to-Neumann map for a nonuniformly elliptic problem posed on a semi-infinite cylinder in one more spatial dimension. This extra dimension further motivates the design of an a posteriori error indicator. The latter is defined as the sum of three contributions, which come from the discretization of the state and adjoint equations and the control variable. The indicator for the state and adjoint equations relies on an anisotropic error estimator in Muckenhoupt weighted Sobolev spaces. The analysis is valid in any dimension. On the basis of the devised a posteriori error estimator, we design a simple adaptive strategy that exhibits optimal experimental rates of convergence.

**Key words.** linear-quadratic optimal control problem, fractional diffusion, nonlocal operators, a posteriori error estimates, anisotropic estimates, adaptive algorithm.

**AMS subject classifications.** 35R11, 35J70, 49J20, 49M25, 65N12, 65N30, 65N50.

**1. Introduction.** In this work we shall be interested in the derivation and analysis of a computable, efficient and, under certain assumptions, reliable a posteriori error estimator for a constrained linear-quadratic optimal control problem involving fractional powers of the Dirichlet Laplace operator. To the best of our knowledge, this is the first work that addresses this problem. To make matters precise, for  $n \geq 1$ , we let  $\Omega$  be an open and bounded polytopal domain of  $\mathbb{R}^n$  with Lipschitz boundary  $\partial\Omega$ . Given  $s \in (0, 1)$ , and a desired state  $u_d : \Omega \rightarrow \mathbb{R}$ , we define the cost functional

$$J(u, z) = \frac{1}{2} \|u - u_d\|_{L^2(\Omega)}^2 + \frac{\mu}{2} \|z\|_{L^2(\Omega)}^2, \quad (1.1)$$

where  $\mu > 0$  is the so-called regularization parameter. With these ingredients at hand, we define the *fractional optimal control problem* as follows: Find

$$\min J(u, z), \quad (1.2)$$

subject to the *fractional state equation*

$$(-\Delta)^s u = z \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega, \quad (1.3)$$

and the *control constraints*

$$a(x') \leq z(x') \leq b(x') \quad \text{a.e. } x' \in \Omega. \quad (1.4)$$

\*The first author's research has been partially supported by the NSF grant DMS-1521590. The second author's research has been partially supported by CONICYT through FONDECYT project 3160201.

<sup>†</sup>Department of Mathematical Sciences, George Mason University, Fairfax, VA 22030, USA. hantil@gmu.edu.

<sup>‡</sup>Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile. enrique.otarola@usm.cl.

The operator  $(-\Delta)^s$ , with  $s \in (0, 1)$ , denotes the fractional powers of the Dirichlet Laplace operator, which for convenience we will simply call the *fractional Laplacian*. The functions  $\mathbf{a}$  and  $\mathbf{b}$  both belong to  $L^2(\Omega)$  and satisfy the property  $\mathbf{a}(x') \leq \mathbf{b}(x')$  for almost every  $x' \in \Omega$ .

A rather incomplete list of problems where fractional derivatives and fractional diffusion appears includes: mechanics [6], where they are used to model viscoelastic behavior [26], turbulence [24, 28] and the hereditary properties of materials [36]; diffusion processes [1, 51], in particular processes in disordered media, where the disorder may change the laws of Brownian motion and thus lead to anomalous diffusion [10, 12]; nonlocal electrostatics [40]; finance [46]; image processing [34]; biophysics [14]; chaotic dynamical systems [59] and many others [13, 27]. Optimal control problems arise naturally in these applications and then it is essential to design numerical schemes to efficiently approximate them.

The analysis of problems involving the fractional Laplacian is delicate and involves fine results in harmonic analysis [45, 61, 62]; one of the main difficulties being the nonlocality of the operator. This difficulty has been resolved to some extent by L. Caffarelli and L. Silvestre [15], who have proposed a technique that turned out to be a breakthrough and has paved the way to study fractional laplacians using *local* techniques. Namely, any power  $s \in (0, 1)$  of the fractional Laplacian in  $\mathbb{R}^n$  can be realized as an operator that maps a Dirichlet boundary condition to a Neumann-type condition via an extension problem on the upper half-space  $\mathbb{R}_+^{n+1}$ . This result was later adapted in [16, 63] to bounded domains  $\Omega$ , thus obtaining an extension problem posed on the semi-infinite cylinder  $\mathcal{C} = \Omega \times (0, \infty)$ . This extension corresponds to the following mixed boundary value problem:

$$\operatorname{div}(y^\alpha \nabla \mathcal{U}) = 0 \text{ in } \mathcal{C}, \quad \mathcal{U} = 0 \text{ on } \partial_L \mathcal{C}, \quad \partial_{\nu^\alpha} \mathcal{U} = d_s \mathbf{z} \quad \text{on } \Omega \times \{0\}, \quad (1.5)$$

where  $\partial_L \mathcal{C} = \partial\Omega \times [0, \infty)$  is the lateral boundary of  $\mathcal{C}$  and  $d_s = 2^\alpha \Gamma(1-s)/\Gamma(s)$  is a positive normalization constant. The parameter  $\alpha$  is defined as  $\alpha = 1 - 2s \in (-1, 1)$  and the conormal exterior derivative of  $\mathcal{U}$  at  $\Omega \times \{0\}$  is

$$\partial_{\nu^\alpha} \mathcal{U} = - \lim_{y \rightarrow 0^+} y^\alpha \mathcal{U}_y. \quad (1.6)$$

We call  $y$  the *extended variable* and call the dimension  $n+1$  in  $\mathbb{R}_+^{n+1}$  the *extended dimension* of problem (1.5). The limit in (1.6) must be understood in the distributional sense; see [15, 16, 63]. With these elements at hand, we then write the fundamental result by L. Caffarelli and L. Silvestre [15, 16, 63]: the fractional Laplacian and the Dirichlet-to-Neumann map of problem (1.5) are related by  $d_s(-\Delta)^s \mathbf{u} = \partial_{\nu^\alpha} \mathcal{U}$  in  $\Omega$ .

The use of the aforementioned localization techniques for the numerical treatment of problem (1.3) followed not so long after [53]. In this reference, the authors propose the following technique to solve problem (1.3): given  $\mathbf{z}$ , solve (1.5), thus obtaining a function  $\mathcal{U}$ ; setting  $\mathbf{u}(x') = \mathcal{U}(x', 0)$ , the solution to (1.3) is obtained. The implementation of this scheme uses standard components of finite element analysis, while its analysis combines asymptotic properties of Bessel functions [2], elements of harmonic analysis [30, 50] and a polynomial interpolation theory on weighted spaces [31, 52]. The latter is valid for tensor product elements that exhibit a large aspect ratio in  $y$  (anisotropy), which is necessary to fit the behavior of  $\mathcal{U}(x', y)$  with  $x' \in \Omega$  and  $y > 0$ . The main advantage of this scheme is that it solves the local problem (1.5) instead of dealing with  $(-\Delta)^s$  in (1.3). However, this comes at the expense of incorporating one more dimension to the problem; issue that has been resolved to some extent with the design of fast solvers [23] and adaptive finite element methods (AFEMs) [22].

Exploiting the ideas developed in [53], in the previous work [4], we have proposed two numerical strategies to approximate the solution to (1.2)–(1.4). Invoking the localization results of [15, 16, 63], we have considered an equivalent optimal control problem:  $\min J(\mathcal{U}(\cdot, 0), \mathbf{z})$  subject to the *linear state equation* (1.5) and the *control constraints* (1.4). Since (1.5) is posed on the semi-infinite cylinder  $\mathcal{C}$ , we have then introduced a truncated optimal control problem and analyzed its approximation properties. On the basis of this, we have proposed two schemes based on the discretization of the state and adjoint equations with first-degree tensor product finite elements on anisotropic meshes: the variational approach [39] and a fully discrete scheme that discretizes the set of controls by piecewise constant functions [5, 18]. The latter yields an optimal error estimate for the control approximation: If  $\Omega$  is convex,  $\mathbf{u}_d \in \mathbb{H}^{1-s}(\Omega)$ , and  $\mathbf{a}, \mathbf{b} \in \mathbb{R}$  are such that  $\mathbf{a} \leq 0 \leq \mathbf{b}$  for  $s \in (0, \frac{1}{2}]$ , then

$$\|\bar{\mathbf{z}} - \bar{Z}\|_{L^2(\Omega)} \lesssim |\log N|^{2s} N^{-\frac{1}{n+1}}, \quad (1.7)$$

where  $\bar{\mathbf{z}}$  denotes the optimal solution to fractional optimal control problem,  $\bar{Z}$  corresponds to the optimal solution of the discrete counterpart of (1.2)–(1.4) and  $N$  denotes the number of the degrees of freedom of the underlying mesh.

Since the aforementioned scheme incorporates one extra dimension, it raises the following question: How efficient is this method? A quest for an answer to this question motivates the study of AFEMs, since it is known that they constitute an efficient class of numerical methods for approximating the solution to optimal control problems [11, 38, 42]: they allow for their resolution with relatively modest computational resources. In addition, they can achieve optimal performance, measured as error versus degrees of freedom, in situations when classical FEM cannot [42, 56, 55]. An essential ingredient of AFEMs is an a posteriori error estimator, which is a computable quantity that depends on the discrete solution and data, and provides information about the local quality of the approximate solution. For linear second-order elliptic boundary value problems, the theory has attained a mature understanding; see [3, 48, 55, 56, 65] for an up-to-date discussion including also the design of AFEMs, their convergence and optimal complexity. In contrast to this well-established theory, the a posteriori error analysis for constrained optimal control problem has not been fully understood yet; the main source of difficulty is its inherent nonlinear feature. We refer the reader to [42] for an up-to-date discussion.

AFEMs for the fractional optimal control problem are also motivated by the fact that the a priori error estimate (1.7) requires  $\bar{\mathbf{z}} \in \mathbb{H}^{1-s}(\Omega)$ , which in turn demands  $\Omega$  convex,  $\mathbf{u}_d \in \mathbb{H}^{1-s}(\Omega)$  and  $\mathbf{a} \leq 0 \leq \mathbf{b}$  for  $s \in (0, \frac{1}{2}]$ . If one of these conditions does not hold, the optimal control  $\bar{\mathbf{z}}$  may have singularities in the  $x'$ -variables and thus exhibits fractional regularity. Consequently, quasi-uniform refinement of  $\Omega$  would not result in an efficient solution technique; see [53, section 6.3] for an illustration of this situation at the level of solving the state equation (1.5).

The main contribution of this work is the design and analysis of a computable, efficient and, under certain assumptions, reliable a posteriori error estimator for the fractional optimal control problem (1.2)–(1.4). As it was highlighted before, there is undoubtedly need for developing such an estimator and this is the first work that addresses this question for problem (1.2)–(1.4). Given a mesh  $\mathcal{T}$  and corresponding approximations  $\mathcal{U}_{\mathcal{T}}$ ,  $\mathcal{P}_{\mathcal{T}}$  and  $\bar{\mathbf{z}}_{\mathcal{T}}$ , the proposed error indicator is built on the basis of three contributions:

$$\mathcal{E}_{\text{ocp}} = \mathcal{E}_{\mathcal{U}} + \mathcal{E}_{\mathcal{P}} + \mathcal{E}_{\mathbf{z}},$$

where  $\mathcal{E}_{\mathcal{U}}$  and  $\mathcal{E}_{\mathcal{P}}$  correspond to the anisotropic posteriori error estimator on weighted Sobolev spaces of [22], for the state and adjoint equations, respectively. The error indicator  $\mathcal{E}_{\mathbf{z}}$  is defined as the  $\ell^2$ -sum of the local contributions  $\mathcal{E}_{\mathbf{z}}(\bar{\mathbf{z}}_{\mathcal{T}}, \bar{\mathcal{P}}_{\mathcal{T}}; T) = \|\bar{\mathbf{z}}_{\mathcal{T}} - \Pi(-\frac{1}{\mu} \bar{\mathcal{P}}_{\mathcal{T}}(\cdot, 0))\|_{L^2(\Omega)}$ , with  $T \in \mathcal{T}$  and  $\Pi(v) = \min\{\mathbf{b}, \max\{\mathbf{a}, v\}\}$ . We present an analysis for  $\mathcal{E}_{\text{ocp}}$ , we prove its efficiency and, under certain assumptions, its reliability. We remark that the devised error estimator is able to deal with both: the natural anisotropy of the mesh  $\mathcal{T}$  in the extended variable and the degenerate coefficient  $y^\alpha$ . This approach is of value not only for the fractional optimal control problem, but in general for control problem involving anisotropic meshes since rigorous anisotropic a posteriori error estimators are scarce in the literature.

**2. Notation and preliminaries.** Throughout this work  $\Omega$  is an open and bounded polytopal domain of  $\mathbb{R}^n$  ( $n \geq 1$ ) with Lipschitz boundary  $\partial\Omega$ . We define the semi-infinite cylinder with base  $\Omega$  and its lateral boundary, respectively, by  $\mathcal{C} = \Omega \times (0, \infty)$  and  $\partial_L \mathcal{C} = \partial\Omega \times [0, \infty)$ . Given  $\mathcal{Y} > 0$ , we define the truncated cylinder  $\mathcal{C}_{\mathcal{Y}} = \Omega \times (0, \mathcal{Y})$  and  $\partial_L \mathcal{C}_{\mathcal{Y}}$  accordingly. If  $x \in \mathbb{R}^{n+1}$ , we write  $x = (x^1, \dots, x^n, x^{n+1}) = (x', x^{n+1}) = (x', y)$ , with  $x^i \in \mathbb{R}$  for  $i = 1, \dots, n+1$ ,  $x' \in \mathbb{R}^n$  and  $y \in \mathbb{R}$ ; this notation distinguishes the extended dimension  $y$ .

We denote by  $(-\Delta)^s$ ,  $s \in (0, 1)$ , a fractional power of Dirichlet Laplace operator  $(-\Delta)$ . The parameter  $\alpha$  belongs to  $(-1, 1)$  and is related to the power  $s$  of the fractional Laplacian  $(-\Delta)^s$  by the formula  $\alpha = 1 - 2s$ .

Finally, the relation  $a \lesssim b$  indicates that  $a \leq Cb$ , with a constant  $C$  that does not depend on  $a$  or  $b$  nor the discretization parameters. The value of  $C$  might change at each occurrence.

**2.1. The fractional Laplace operator.** We adopt the spectral definition for the fractional powers of the Dirichlet Laplace operator [16, 53]. The operator  $(-\Delta)^{-1} : L^2(\Omega) \rightarrow L^2(\Omega)$  that solves  $-\Delta w = f$  in  $\Omega$  and  $w = 0$  on  $\partial\Omega$ , is compact, symmetric and positive, whence its spectrum  $\{\lambda_k^{-1}\}_{k \in \mathbb{N}}$  is discrete, real, positive and accumulates at zero. Moreover, the eigenfunctions

$$\{\varphi_k\}_{k \in \mathbb{N}} : \quad -\Delta \varphi_k = \lambda_k \varphi_k \text{ in } \Omega, \quad \varphi_k = 0 \text{ on } \Omega, \quad k \in \mathbb{N},$$

form an orthonormal basis of  $L^2(\Omega)$ . Fractional powers of  $(-\Delta)$  can be defined by

$$(-\Delta)^s w := \sum_{k=1}^{\infty} \lambda_k^s w_k \varphi_k, \quad w \in C_0^\infty(\Omega), \quad s \in (0, 1),$$

where  $w_k = \int_{\Omega} w \varphi_k$ . By density we extend this definition to

$$\mathbb{H}^s(\Omega) = \left\{ w = \sum_{k=1}^{\infty} w_k \varphi_k : \sum_{k=1}^{\infty} \lambda_k^s w_k^2 < \infty \right\} = [H_0^1(\Omega), L^2(\Omega)]_{1-s};$$

see [53] for details. For  $s \in (0, 1)$  we denote by  $\mathbb{H}^{-s}(\Omega)$  the dual space of  $\mathbb{H}^s(\Omega)$ .

**2.2. The Caffarelli–Silvestre extension problem.** In this section we explore problem (1.5) and its relation with the nonlocal problem (1.3); we refer the reader to [15, 16, 53, 63] for details. Since  $\alpha \in (-1, 1)$ , problem (1.5) is nonuniformly elliptic and thus it requires to introduce weighted Lebesgue and Sobolev spaces for its description. Let  $E$  be an open set in  $\mathbb{R}^{n+1}$ . We define  $L^2(|y|^\alpha, E)$  as the Lebesgue space for the

measure  $|y|^\alpha dx$ . We also define the weighted Sobolev space  $H^1(|y|^\alpha, E) := \{w \in L^2(|y|^\alpha, E) : |\nabla w| \in L^2(|y|^\alpha, E)\}$ , which we endow with the norm

$$\|w\|_{H^1(|y|^\alpha, E)} = \left( \|w\|_{L^2(|y|^\alpha, E)}^2 + \|\nabla w\|_{L^2(|y|^\alpha, E)}^2 \right)^{\frac{1}{2}}. \quad (2.1)$$

Since  $\alpha = 1 - 2s \in (-1, 1)$ , the weight  $|y|^\alpha$  belongs to the Muckenhoupt class  $A_2(\mathbb{R}^{n+1})$  [30, 64]. Consequently,  $H^1(|y|^\alpha, D)$  is Hilbert and  $C^\infty(\mathcal{D}) \cap H^1(|y|^\alpha, D)$  is dense in  $H^1(|y|^\alpha, D)$  (cf. [64, Proposition 2.1.2, Corollary 2.1.6] and [35, Theorem 1]).

The natural space to seek for a weak solution to problem (1.5) is

$$\mathring{H}_L^1(y^\alpha, \mathcal{C}) := \{w \in H^1(y^\alpha, \mathcal{C}) : w = 0 \text{ on } \partial_L \mathcal{C}\}.$$

We recall the following *weighted Poincaré inequality* [53, inequality (2.21)]:

$$\|w\|_{L^2(y^\alpha, \mathcal{C})} \lesssim \|\nabla w\|_{L^2(y^\alpha, \mathcal{C})} \quad \forall w \in \mathring{H}_L^1(y^\alpha, \mathcal{C}).$$

This yields that the seminorm on  $\mathring{H}_L^1(y^\alpha, \mathcal{C})$  is equivalent to (2.1). For  $w \in H^1(y^\alpha, \mathcal{C})$   $\text{tr}_\Omega w$  denotes its trace onto  $\Omega \times \{0\}$ . We recall ([53, Prop. 2.5] and [16, Prop. 2.1])

$$\text{tr}_\Omega \mathring{H}_L^1(y^\alpha, \mathcal{C}) = \mathbb{H}^s(\Omega), \quad \|\text{tr}_\Omega w\|_{\mathbb{H}^s(\Omega)} \leq C_{\text{tr}_\Omega} \|w\|_{\mathring{H}_L^1(y^\alpha, \mathcal{C})}. \quad (2.2)$$

We must mention that  $C_{\text{tr}_\Omega} \leq d_s^{-\frac{1}{2}}$  [22, section 2.3], where  $d_s = 2^\alpha \Gamma(1-s)/\Gamma(s)$ . This will be useful in the analysis of the proposed a posteriori error indicator.

We conclude this section with the fundamental result by Caffarelli and Silvestre [15, 16, 63]: If  $u \in \mathbb{H}^s(\Omega)$  and  $\mathcal{U} \in \mathring{H}_L^1(y^\alpha, \mathcal{C})$  solve (1.3) and (1.5), respectively, then

$$d_s(-\Delta)^s u = \partial_{\nu^\alpha} \mathcal{U} = - \lim_{y \rightarrow 0^+} y^\alpha \mathcal{U}_y,$$

in the sense of distributions. Here,  $s \in (0, 1)$  and  $\alpha = 1 - 2s \in (-1, 1)$ .

**3. A priori error estimates.** In an effort to make this work self-contained, in this section we review the results of [4], where an a priori error analysis for a fully discrete approximation of the fractional optimal control problem is investigated. This will also serve to make clear the limitations of this theory.

**3.1. The extended optimal control problem.** We start by recalling an equivalent problem to (1.2)–(1.4): the *extended optimal control problem*. The main advantage of this problem is its local nature and is based on the Caffarelli–Silvestre extension result. To describe it, we define the set of *admissible controls* as

$$\mathbf{Z}_{\text{ad}} = \{w \in L^2(\Omega) : a(x') \leq w(x') \leq b(x') \text{ a.e. } x' \in \Omega\}, \quad (3.1)$$

where  $a, b \in L^2(\Omega)$  and satisfy the property  $a(x') \leq b(x')$  a.e.  $x' \in \Omega$ . The extended optimal control problem is then defined as follows: Find  $\min J(\text{tr}_\Omega \mathcal{U}, z)$ , subject to the *linear* state equation

$$a(\mathcal{U}, \phi) = \langle z, \text{tr}_\Omega \phi \rangle \quad \forall \phi \in \mathring{H}_L^1(y^\alpha, \mathcal{C}), \quad (3.2)$$

and the control constraints  $z \in \mathbf{Z}_{\text{ad}}$ . The functional  $J$  is defined by (1.2) with  $u_d \in L^2(\Omega)$  and  $\mu > 0$ . For  $w, \phi \in \mathring{H}_L^1(y^\alpha, \mathcal{C})$ , the bilinear form  $a$  is defined by

$$a(w, \phi) = \frac{1}{d_s} \int_{\mathcal{C}} y^\alpha \nabla w \cdot \nabla \phi$$

and  $\langle \cdot, \cdot \rangle$  denotes the duality pairing between  $\mathbb{H}^s(\Omega)$  and  $\mathbb{H}^{-s}(\Omega)$  which, as a consequence of (2.2), is well defined for  $\mathbf{z} \in \mathbb{H}^{-s}(\Omega)$  and  $\phi \in \dot{H}_L^1(y^\alpha, \mathcal{C})$ .

The extended optimal control problem has a unique optimal solution  $(\bar{\mathcal{U}}, \bar{\mathbf{z}}) \in \dot{H}_L^1(y^\alpha, \mathcal{C}) \times \mathbb{H}^s(\Omega)$  [4, Theorem 3.11] and is equivalent to the fractional optimal control problem:  $\text{tr}_\Omega \bar{\mathcal{U}} = \bar{\mathbf{u}}$  [4, Theorem 3.12].

**3.2. The truncated optimal control problem.** Since  $\mathcal{C}$  is unbounded, problem (3.2) cannot be directly approximated with finite-element-like techniques. However, as [53, Proposition 3.1] shows, the solution  $\mathcal{U}$  of problem (3.2) decays exponentially in the extended variable  $y$ . This suggests to consider a *truncated optimal control problem*, which is based on a truncation of the state equation (3.2). To describe it, we define

$$\dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma) = \{w \in H^1(y^\alpha, \mathcal{C}_\gamma) : w = 0 \text{ on } \partial_L \mathcal{C}_\gamma \cup \Omega \times \{\mathcal{Y}\}\},$$

and for all  $w, \phi \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma)$ , the bilinear form

$$a_\gamma(w, \phi) = \frac{1}{d_s} \int_{\mathcal{C}_\gamma} y^\alpha \nabla w \cdot \nabla \phi. \quad (3.3)$$

The truncated optimal control problem is then defined as follows: Find  $\min J(\text{tr}_\Omega v, \mathbf{r})$  subject to the truncated state equation

$$a_\gamma(v, \phi) = \langle \mathbf{r}, \text{tr}_\Omega \phi \rangle \quad \forall \phi \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma) \quad (3.4)$$

and the control constraints  $\mathbf{r} \in \mathbf{Z}_{\text{ad}}$ . The existence and uniqueness of an optimal pair  $(\bar{v}, \bar{\mathbf{r}}) \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma) \times \mathbb{H}^s(\Omega)$  follows from [4, Theorem 4.5]. In addition, we have that the optimal control  $\bar{\mathbf{r}} \in \mathbf{Z}_{\text{ad}}$  verifies the variational inequality

$$(\text{tr}_\Omega \bar{p} + \mu \bar{\mathbf{r}}, \mathbf{r} - \bar{\mathbf{r}})_{L^2(\Omega)} \geq 0 \quad \forall \mathbf{r} \in \mathbf{Z}_{\text{ad}}, \quad (3.5)$$

where  $\bar{p} \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma)$  denotes the optimal adjoint state and solves

$$a_\gamma(\bar{p}, \phi) = (\text{tr}_\Omega \bar{v} - \mathbf{u}_d, \text{tr}_\Omega \phi)_{L^2(\Omega)}. \quad (3.6)$$

The following approximation properties follow from [4, Lemma 4.6]: If  $(\bar{\mathcal{U}}, \bar{\mathbf{z}}) \in \dot{H}_L^1(y^\alpha, \mathcal{C}) \times \mathbb{H}^s(\Omega)$  and  $(\bar{v}, \bar{\mathbf{r}}) \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma) \times \mathbb{H}^s(\Omega)$  solve the extended and truncated optimal control problems, respectively, then

$$\begin{aligned} \|\bar{\mathbf{z}} - \bar{\mathbf{r}}\|_{L^2(\Omega)} &\lesssim e^{-\sqrt{\lambda_1} \mathcal{Y}/4} (\|\bar{\mathbf{r}}\|_{L^2(\Omega)} + \|\mathbf{u}_d\|_{L^2(\Omega)}), \\ \|\nabla(\mathcal{U} - \bar{v})\|_{L^2(y^\alpha, \mathcal{C})} &\lesssim e^{-\sqrt{\lambda_1} \mathcal{Y}/4} (\|\bar{\mathbf{r}}\|_{L^2(\Omega)} + \|\mathbf{u}_d\|_{L^2(\Omega)}), \end{aligned}$$

where  $\lambda_1$  denotes the first eigenvalue of the operator  $-\Delta$ .

**3.3. A fully discrete scheme.** In this section we recall the fully discrete scheme, proposed in [4, section 5.3], that approximates the solution to (1.2)–(1.4). We also review its a priori error analysis; see [4, section 5.3] for details. To do so in this section, and this section only, we will assume the following regularity result, which is valid if, for instance, the domain  $\Omega$  is convex [37]:

$$\|w\|_{H^2(\Omega)} \lesssim \|\Delta_{x'} w\|_{L^2(\Omega)} \quad \forall w \in H^2(\Omega) \cap H_0^1(\Omega). \quad (3.7)$$

The analysis of the fully discrete scheme of [4, section 5.3] relies on the regularity properties of the optimal pairs  $(\bar{\mathcal{U}}, \bar{\mathbf{z}})$  and  $(\bar{v}, \bar{\mathbf{r}})$  that solve the extended and truncated

optimal control problems, respectively. We review such regularity properties in what follows. The results of [53, Theorem 2.7] reveals that the second order regularity of  $\mathcal{U}$ , solving (3.2), is much worse in the extended direction, namely

$$\|\Delta_{x'} \mathcal{U}\|_{L^2(y^\alpha, \mathcal{C})} + \|\partial_y \nabla_{x'} \mathcal{U}\|_{L^2(y^\alpha, \mathcal{C})} \lesssim \|\mathbf{z}\|_{\mathbb{H}^{1-s}(\Omega)}, \quad (3.8)$$

$$\|\mathcal{U}_{yy}\|_{L^2(y^\beta, \mathcal{C})} \lesssim \|\mathbf{z}\|_{L^2(\Omega)}, \quad (3.9)$$

where  $\beta > 2\alpha + 1$ . These result are also valid for the solution  $v$  of problem (3.4); see [54, Remark 4.4].

The estimates (3.8)–(3.9) have important consequences in the design of efficient numerical techniques to solve (3.2); they suggest that a graded mesh in the extended  $(n+1)$ -dimension must be used [53, section 5]. We recall the construction of the mesh over  $\mathcal{C}_\gamma$  used in [4, 53]. First, we consider a graded partition  $\mathcal{I}_\gamma$  of the interval  $[0, \gamma]$  with mesh points

$$y_k = \left(\frac{k}{M}\right)^\gamma \gamma, \quad k = 0, \dots, M, \quad (3.10)$$

and  $\gamma > 3/(1-\alpha) = 3/(2s) > 1$ . Second, we consider  $\mathcal{T}_\Omega = \{K\}$  to be a conforming mesh of  $\Omega$ , where  $K \subset \mathbb{R}^n$  is an element that is isoparametrically equivalent either to the unit cube  $[0, 1]^n$  or the unit simplex in  $\mathbb{R}^n$ . We denote by  $\mathbb{T}_\Omega$  the collections of all conforming refinements of an original mesh  $\mathcal{T}_\Omega^0$ . We assume that  $\mathbb{T}_\Omega$  is shape regular [25]. We then construct a mesh  $\mathcal{T}_\gamma$  over  $\mathcal{C}_\gamma$  as the tensor product triangulation of  $\mathcal{T}_\Omega \in \mathbb{T}_\Omega$  and  $\mathcal{I}_\gamma$ . We denote by  $\mathbb{T}$  the set of all the meshes obtained with this procedure, and recall that  $\mathbb{T}$  satisfies the following weak shape regularity condition: If  $T_1 = K_1 \times I_1$  and  $T_2 = K_2 \times I_2 \in \mathcal{T}_\gamma$  have nonempty intersection, then there exists a positive constant  $\sigma_\gamma$  such that

$$h_{I_1} h_{I_2}^{-1} \leq \sigma_\gamma, \quad (3.11)$$

where  $h_I = |I|$ . This weak shape regularity condition allows for anisotropy in the extended variable  $y$  [31, 53, 52].

For  $\mathcal{T}_\gamma \in \mathbb{T}$ , we define the finite element space

$$\mathbb{V}(\mathcal{T}_\gamma) = \{W \in C^0(\overline{\mathcal{C}_\gamma}) : W|_T \in \mathcal{P}_1(K) \otimes \mathbb{P}_1(I) \ \forall T \in \mathcal{T}_\gamma, \ W|_{\Gamma_D} = 0\}, \quad (3.12)$$

where  $\Gamma_D = \partial_L \mathcal{C}_\gamma \cup \Omega \times \{\gamma\}$  is the Dirichlet boundary. The space  $\mathcal{P}_1(K)$  is  $\mathbb{P}_1(K)$  – the space of polynomials of degree at most 1, when the base  $K$  of  $T = K \times I$  is a simplex. If  $K$  is a cube,  $\mathcal{P}_1(K)$  stand for  $\mathbb{Q}_1(K)$  – the space of polynomials of degree not larger than 1 in each variable. We also define the space  $\mathbb{U}(\mathcal{T}_\Omega) = \text{tr}_\Omega \mathbb{V}(\mathcal{T}_\gamma)$ , which is simply a  $\mathcal{P}_1$  finite element space over the mesh  $\mathcal{T}_\Omega$ .

Before describing the numerical scheme introduced and developed in [4], we recall the regularity properties of the extended and truncated optimal controls  $\bar{\mathbf{z}}$  and  $\bar{\mathbf{r}}$ , respectively. If  $\mathbf{u}_d \in \mathbb{H}^{1-s}(\Omega)$  and  $\mathbf{a} \leq 0 \leq \mathbf{b}$  for  $s \in (0, \frac{1}{2}]$ , then  $\bar{\mathbf{z}} \in H^1(\Omega) \cap \mathbb{H}^{1-s}(\Omega)$  [4, Lemmas 3.5 and 5.9]. Under the same framework, we have the same result for the truncated optimal control:  $\bar{\mathbf{r}} \in H^1(\Omega) \cap \mathbb{H}^{1-s}(\Omega)$  [4, Lemma 5.9].

After all these preparations, we are ready to describe the fully discrete scheme to approximate the fractional optimal control problem. The *fully discrete optimal control problem* reads as follows:  $\min J(\text{tr}_\Omega V, Z)$ , subject to the discrete state equation

$$a_\gamma(V, W) = (Z, \text{tr}_\Omega W)_{L^2(\Omega)} \quad \forall W \in \mathbb{V}(\mathcal{T}_\gamma), \quad (3.13)$$

and the discrete control constraints  $Z \in \mathbb{Z}_{ad}(\mathcal{T}_\Omega)$ . We recall that the functional  $J$ , the bilinear form  $a_{\mathcal{Y}}$  and the discrete space  $\mathbb{V}(\mathcal{T}_{\mathcal{Y}})$  are defined by (1.1), (3.3), and (3.12), respectively. The discrete and admissible set of controls is defined by

$$\mathbb{Z}_{ad}(\mathcal{T}_\Omega) = \mathbb{Z}_{ad} \cap \{Z \in L^\infty(\Omega) : Z|_K \in \mathbb{P}_0(K) \quad \forall K \in \mathcal{T}_\Omega\},$$

i.e., the space of piecewise constant functions defined on the partition  $\mathcal{T}_\Omega$  that verifies the control bounds, which we assume to be real constants.

The existence and uniqueness of an optimal pair  $(\bar{V}, \bar{Z}) \in \mathbb{V}(\mathcal{T}_{\mathcal{Y}}) \times \mathbb{Z}_{ad}(\mathcal{T}_\Omega)$  solving the aforementioned problem is standard [4, Theorem 5.15]. In addition, the optimal control  $\bar{Z} \in \mathbb{Z}_{ad}(\mathcal{T}_\Omega)$  is uniquely characterized by the variational inequality

$$(\text{tr}_\Omega \bar{P} + \mu \bar{Z}, Z - \bar{Z})_{L^2(\Omega)} \geq 0 \quad \forall Z \in \mathbb{Z}_{ad}(\mathcal{T}_\Omega), \quad (3.14)$$

where the optimal and discrete adjoint state  $\bar{P} \in \mathbb{V}(\mathcal{T}_{\mathcal{Y}})$  solves

$$a_{\mathcal{Y}}(\bar{P}, W) = (\text{tr}_\Omega \bar{V} - u_d, \text{tr}_\Omega W)_{L^2(\Omega)} \quad \forall W \in \mathbb{V}(\mathcal{T}_{\mathcal{Y}}). \quad (3.15)$$

With the discrete solution  $\bar{V} \in \mathbb{V}(\mathcal{T}_{\mathcal{Y}})$  at hand, we define

$$\bar{U} := \text{tr}_\Omega \bar{V}, \quad (3.16)$$

and thus obtain a fully discrete approximation  $(\bar{U}, \bar{Z}) \in \mathbb{U}(\mathcal{T}_\Omega) \times \mathbb{Z}_{ad}(\mathcal{T}_\Omega)$  of the optimal pair  $(\bar{u}, \bar{z}) \in \mathbb{H}^s(\Omega) \times \mathbb{Z}_{ad}$  solving the fractional optimal control problem.

To write the a priori error estimates for the fully discrete optimal control problem, we notice that  $\#\mathcal{T}_{\mathcal{Y}} = M \#\mathcal{T}_\Omega$ , and that  $\#\mathcal{T}_\Omega \approx M^n$  implies  $\#\mathcal{T}_{\mathcal{Y}} \approx M^{n+1}$ . Consequently, if  $\mathcal{T}_\Omega$  is quasi-uniform, we have that  $h_{\mathcal{T}_\Omega} \approx (\#\mathcal{T}_\Omega)^{-1/n}$ . We then have the following result [4, Corollary 5.17].

**THEOREM 3.1** (fractional control problem: error estimate). *Let  $(\bar{V}, \bar{Z}) \in \mathbb{V}(\mathcal{T}_{\mathcal{Y}}) \times \mathbb{Z}_{ad}$  solves the fully discrete control problem and  $\bar{U} \in \mathbb{U}(\mathcal{T}_\Omega)$  be defined as in (3.16). If  $\Omega$  verifies (3.7),  $u_d \in \mathbb{H}^{1-s}(\Omega)$ , and  $\mathbf{a} \leq 0 \leq \mathbf{b}$  for  $s \in (0, \frac{1}{2}]$ , then we have*

$$\|\bar{z} - \bar{Z}\|_{L^2(\Omega)} \lesssim |\log(\#\mathcal{T}_{\mathcal{Y}})|^{2s} (\#\mathcal{T}_{\mathcal{Y}})^{\frac{-1}{n+1}} (\|\bar{r}\|_{H^1(\Omega)} + \|u_d\|_{\mathbb{H}^{1-s}(\Omega)}), \quad (3.17)$$

and

$$\|\bar{u} - \bar{U}\|_{\mathbb{H}^s(\Omega)} \lesssim |\log(\#\mathcal{T}_{\mathcal{Y}})|^{2s} (\#\mathcal{T}_{\mathcal{Y}})^{\frac{-1}{n+1}} (\|\bar{r}\|_{H^1(\Omega)} + \|u_d\|_{\mathbb{H}^{1-s}(\Omega)}), \quad (3.18)$$

where the truncation parameter  $\mathcal{Y}$ , in the truncated optimal control problem, is chosen such that  $\mathcal{Y} \approx \log(\#\mathcal{T}_{\mathcal{Y}})$ .

**REMARK 3.2** (Domain and data regularity). The results of Theorem 3.1 are valid if and only if  $\Omega$  is such that (3.7) holds,  $u_d \in \mathbb{H}^{1-s}(\Omega)$ , and  $\mathbf{a} \leq 0 \leq \mathbf{b}$  for  $s \in (0, \frac{1}{2}]$ .

**4. A posteriori error analysis.** The design and analysis of a posteriori error estimators for linear second-order elliptic boundary value problems on isotropic discretizations, i.e., meshes where the aspect ratio of all cells is bounded independently of the refinement level, has achieved a certain degree of maturity. Starting with the pioneering work of Babuška and Rheinboldt [8], a great deal of work has been devoted to its study. We refer the reader to [3, 9, 49, 55, 56, 65] for an up-to-date discussion including also the design of AFEMs, their convergence and optimal complexity. In contrast to this well-established theory, the a posteriori error estimation on anisotropic discretizations, i.e., meshes where the cells have disparate sizes in each direction, is



still not completely understood. To the best of our knowledge, the first work that introduces an a posteriori error estimator on anisotropic meshes is [60]. The analysis provided in this work relies on certain assumptions on the mesh [60, section 2], on the exact solution [60, Definition 3.1], and on the discrete solution [60, Definition 5.2]. However, no explicit examples of AFEMs satisfying these assumptions are provided and their construction is not evident. Afterwards, the so-called *matching function* is introduced in [43, 44] for deriving error indicators on anisotropic meshes. The presented analysis relies on the correct alignment of the grid with the exact solution. Indeed, the upper bound for the error involves the matching function, which depends on the error itself and then it does not provide a real computable quantity; see [43, Theorem 2] and see [44, Theorem 5.1]. The effect of approximating the matching function with a recovered gradient based technique is discussed in [43, 44].

To the best of our knowledge, the first paper that attempts to deal with an anisotropic a posteriori error estimator for an optimal control problem is [57]. In this work, the author proposes, based on the goal-oriented approach developed in [11], an anisotropic error indicator for a parabolic optimal control problem involving the heat equation. However, the presented upper bound for the error [57, Proposition 7] depends on the exact solution and therefore, it is not computable; see the discussion in [57, section 5]. Later, reference [47] presents an anisotropic posteriori error estimator for an optimal control problem of a scalar advection–reaction–diffusion equation. The analysis relies on the goal-oriented approach of [11], and the a priori and posteriori error analyses of [32] and [33], respectively. The presented upper bound for the error depends on the exact optimal variables and therefore is not computable [47, Proposition 3.5]. This shortcoming is circumvented, computationally, by invoking a suitable recovery procedure.

The main contribution of this work is the design and study of an a posteriori error indicator for the fractional optimal control problem (1.2)–(1.4). To accomplish this task, we invoke the a posteriori error indicator developed in [22] that is based on the solution of local problems on stars; we remark that, since problems (3.4) and (3.6) involve the coefficient  $y^\alpha$  ( $-1 < \alpha < 1$ ), that is not uniformly bounded, the usual residual estimator does not apply. The idea of working on stars goes back to Babuška and Miller [7], who introduced local Dirichlet problems. Later, references [17, 49] proposed solving local weighted problems on stars that deliver rather good effectivity indices. A convergence proof of AFEM driven by such error indicators is provided in [49] for a Poisson problem, and in [20] for a general second-order elliptic PDE; the latter also includes optimal complexity. We also refer the reader to [58] for estimators based on solving Neumann problems on elements and their further improvements via the so-called *flux equilibration principle* [3].

Concerning the a posteriori error analysis for (1.2)–(1.4), we first propose and explore an ideal anisotropic error indicator that is constructed on the basis of solving local problems on *cylindrical stars*. This indicator is able to deal with both: the coefficient  $y^\alpha$  and the anisotropic mesh  $\mathcal{T}_\gamma$ . Under a computationally implementable geometric condition imposed on the mesh, which does not depend on the exact optimal variables, we derive the equivalence between the ideal estimator and the error without oscillation terms. This ideal indicator sets the basis to define a computable error estimator, which, under certain assumptions, is equivalent to the error up to data oscillations terms.

**4.1. Preliminaries.** Let us begin the discussion on a posteriori error estimation with some terminology and notation that follows from [22]. Given a node  $z$  on the

mesh  $\mathcal{T}_{\mathcal{Y}}$ , we write  $z = (z', z'')$  where  $z'$  and  $z''$  are nodes on the meshes  $\mathcal{T}_{\Omega}$  and  $\mathcal{T}_{\mathcal{Y}}$  respectively.

Given  $K \in \mathcal{T}_{\Omega}$ , we denote by  $\mathcal{N}(K)$  the set of nodes of  $K$  and by  $\mathring{\mathcal{N}}(K)$  the set of interior nodes. With this notation at hand, we define  $\mathcal{N}(\mathcal{T}_{\Omega}) = \cup\{\mathcal{N}(K) : K \in \mathcal{T}_{\Omega}\}$  and  $\mathring{\mathcal{N}}(\mathcal{T}_{\Omega}) = \cup\{\mathring{\mathcal{N}}(K) : K \in \mathcal{T}_{\Omega}\}$ . Given  $T \in \mathcal{T}_{\mathcal{Y}}$ , we define  $\mathcal{N}(T)$ ,  $\mathring{\mathcal{N}}(T)$ , and then  $\mathring{\mathcal{N}}(\mathcal{T}_{\mathcal{Y}})$  and  $\mathcal{N}(\mathcal{T}_{\mathcal{Y}})$  accordingly.

Given  $z' \in \mathcal{N}(\mathcal{T}_{\Omega})$ , we define the *star* around  $z'$  as

$$S_{z'} = \bigcup \{K \in \mathcal{T}_{\Omega} : K \ni z'\} \subset \Omega$$

and the *cylindrical star* around  $z'$  as

$$\mathcal{C}_{z'} := \bigcup \{T \in \mathcal{T}_{\mathcal{Y}} : T = K \times I, K \ni z'\} = S_{z'} \times (0, \mathcal{Y}) \subset \mathcal{C}_{\mathcal{Y}}. \quad (4.1)$$

Given  $K \in \mathcal{T}_{\Omega}$  we define its *patch* as  $S_K := \bigcup_{z' \in K} S_{z'}$ . For  $T \in \mathcal{T}_{\mathcal{Y}}$  its patch  $S_T$  is defined similarly. Given  $z' \in \mathcal{N}(\mathcal{T}_{\Omega})$  we define its *cylindrical patch* as

$$\mathcal{D}_{z'} := \bigcup \{\mathcal{C}_{w'} : w' \in S_{z'}\} \subset \mathcal{C}_{\mathcal{Y}}.$$

For each  $z' \in \mathcal{N}(\mathcal{T}_{\Omega})$  we set  $h_{z'} := \min\{h_K : K \ni z'\}$ .

**4.2. Local weighted Sobolev spaces.** To define the a posteriori error estimator proposed in this work, we need to introduce some local weighted Sobolev spaces.

DEFINITION 4.1 (local spaces). *Given  $z' \in \mathcal{N}(\mathcal{T}_{\Omega})$ , we define*

$$\mathbb{W}(\mathcal{C}_{z'}) = \{w \in H^1(y^{\alpha}, \mathcal{C}_{z'}) : w = 0 \text{ on } \partial\mathcal{C}_{z'} \setminus \Omega \times \{0\}\}, \quad (4.2)$$

where  $\mathcal{C}_{z'}$  denotes the cylindrical star around  $z'$  defined in (4.1).

Since  $y^{\alpha}$  belongs to the class  $A_2(\mathbb{R}^{n+1})$  [30, 50], the space  $\mathbb{W}(\mathcal{C}_{z'})$  is Hilbert. In addition, we have the following weighted Poincaré-type inequality [22, Proposition 5.8]: If  $w \in \mathbb{W}(\mathcal{C}_{z'})$ , then

$$\|w\|_{L^2(y^{\alpha}, \mathcal{C}_{z'})} \lesssim \mathcal{Y} \|\nabla w\|_{L^2(y^{\alpha}, \mathcal{C}_{z'})}, \quad (4.3)$$

where  $\mathcal{Y}$  denotes the truncation parameter introduced in section 3.2. We also have the following trace inequality that follows from [16, Proposition 2.1]: If  $w \in \mathbb{W}(\mathcal{C}_{z'})$ , then

$$\|\text{tr}_{\Omega} w\|_{L^2(S_{z'})} \leq C_{\text{tr}_{\Omega}} \|\nabla w\|_{L^2(y^{\alpha}, \mathcal{C}_{z'})}. \quad (4.4)$$

We notice that the same arguments of [22, Section 2.3] yield  $C_{\text{tr}_{\Omega}} \leq d_s^{-\frac{1}{2}}$ .

**4.3. An ideal a posteriori error estimator.** On the basis of the notation introduced in subsections 4.1 and 4.2, we propose and analyze an ideal a posteriori error estimator for the fractional optimal control problem (1.2)–(1.4). The proposed error indicator is ideal because it is not computable: it is based on the resolution of local problems on infinite dimensional spaces. However, it provides the intuition required to define a discrete and computable error indicator, as is explained in section 4.4. The construction of this ideal indicator allows for the anisotropic meshes  $\mathcal{T}_{\mathcal{Y}}$  defined in section 3 and the nonuniformly coefficient  $y^{\alpha}$  of problem (3.2). We prove that is equivalent to the error without oscillation terms.

The ideal error indicator is defined as the sum of three contributions:

$$\mathcal{E}_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{T}_\gamma) = \mathcal{E}_V(\bar{V}, \bar{Z}; \mathcal{N}(\mathcal{T}_\Omega)) + \mathcal{E}_P(\bar{P}, \bar{V}; \mathcal{N}(\mathcal{T}_\Omega)) + \mathcal{E}_Z(\bar{Z}, \bar{P}; \mathcal{T}_\Omega), \quad (4.5)$$

where  $\mathcal{T}_\gamma \in \mathbb{T}$  corresponds to the anisotropic mesh constructed in subsection 3.3 and  $\bar{V}$ ,  $\bar{P}$  and  $\bar{Z}$  denote the optimal variables solving the fully discrete optimal control problem described in subsection 3.3. We now proceed to describe each contribution in (4.5) separately. To do this, we introduce, for  $w, \psi \in \mathbb{W}(\mathcal{C}_{z'})$ , the bilinear form

$$a_{z'}(w, \psi) = \frac{1}{d_s} \int_{\mathcal{C}_{z'}} y^\alpha \nabla w \nabla \psi. \quad (4.6)$$

Then, the first contribution in (4.5) is defined on the basis of the indicator developed in [22, section 5.3]. We define  $\zeta_{z'} \in \mathbb{W}(\mathcal{C}_{z'})$  as the solution to

$$a_{z'}(\zeta_{z'}, \psi) = \langle \bar{Z}, \text{tr}_\Omega \psi \rangle - a_{z'}(\bar{V}, \psi) \quad \forall \psi \in \mathbb{W}(\mathcal{C}_{z'}), \quad (4.7)$$

where we recall that the space  $\mathbb{W}(\mathcal{C}_{z'})$  is defined in (4.2). With this definition at hand, we then define the local error estimator

$$\mathcal{E}_V(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) := \|\nabla \zeta_{z'}\|_{L^2(y^\alpha, \mathcal{C}_{z'})} \quad (4.8)$$

and the global error estimator  $\mathcal{E}_V(\bar{V}, \bar{Z}; \mathcal{N}(\mathcal{T}_\Omega)) := \left( \sum_{z' \in \mathcal{N}(\mathcal{T}_\Omega)} \mathcal{E}_V^2(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) \right)^{\frac{1}{2}}$ .

We now describe the second contribution in (4.5). To accomplish this task, we define  $\chi_{z'} \in \mathbb{W}(\mathcal{C}_{z'})$  as the solution to the local problem

$$a_{z'}(\chi_{z'}, \psi) = \langle \text{tr}_\Omega \bar{V} - u_d, \text{tr}_\Omega \psi \rangle - a_{z'}(\bar{P}, \psi) \quad \forall \psi \in \mathbb{W}(\mathcal{C}_{z'}). \quad (4.9)$$

We then define the local error indicator

$$\mathcal{E}_P(\bar{P}, \bar{V}; \mathcal{C}_{z'}) := \|\nabla \chi_{z'}\|_{L^2(y^\alpha, \mathcal{C}_{z'})} \quad (4.10)$$

and the global error indicator  $\mathcal{E}_P(\bar{P}, \bar{V}; \mathcal{N}(\mathcal{T}_\Omega)) := \left( \sum_{z' \in \mathcal{N}(\mathcal{T}_\Omega)} \mathcal{E}_P^2(\bar{P}, \bar{V}; \mathcal{C}_{z'}) \right)^{\frac{1}{2}}$ .

Finally, we define a global error estimator for the optimal control as follows:

$$\mathcal{E}_Z(\bar{Z}, \bar{P}; \mathcal{T}_\Omega) := \left( \sum_{K \in \mathcal{T}_\Omega} \mathcal{E}_Z^2(\bar{Z}, \bar{P}; K) \right)^{1/2}, \quad (4.11)$$

with the local error indicators

$$\mathcal{E}_Z(\bar{Z}, \bar{P}; K) := \|\bar{Z} - \Pi(-\frac{1}{\mu} \text{tr}_\Omega \bar{P})\|_{L^2(K)}. \quad (4.12)$$

In (4.12),  $\Pi : L^2(\Omega) \rightarrow \mathbf{Z}_{\text{ad}}$  denotes the nonlinear projection operator defined by

$$\Pi(x') = \min\{\mathbf{b}, \max\{\mathbf{a}, x'\}\}, \quad (4.13)$$

where  $\mathbf{a}$  and  $\mathbf{b}$  denote the control bounds defining the set  $\mathbf{Z}_{\text{ad}}$  in (3.1).

To invoke the results of [22, section 5.3], we introduce an implementable geometric condition that will allow us to consider graded meshes in  $\Omega$  while preserving the anisotropy in the extended direction  $y$  that is necessary to retain optimal orders of

approximation. The flexibility of having graded meshes in  $\Omega$  is essential for compensating some possible singularities in the  $x'$ -variables. We thus assume the following condition over the family of triangulations  $\mathbb{T}$ : there exists a positive constant  $C_{\mathbb{T}}$  such that, for every mesh  $\mathcal{T}_{\mathcal{Y}} \in \mathbb{T}$ , we have that

$$h_{\mathcal{Y}} \leq C_{\mathbb{T}} h_{z'}, \quad (4.14)$$

for all interior nodes  $z'$  of  $\mathcal{T}_{\Omega}$ . Here,  $h_{\mathcal{Y}}$  denotes the largest size in the  $y$ -direction. We remark that this condition is fully implementable.

We now derive an estimate of the energy error in terms of the total error estimator  $\mathcal{E}_{\text{ocp}}$  defined in (4.5) (reliability).

**THEOREM 4.2** (global upper bound). *Let  $(\bar{v}, \bar{p}, \bar{r}) \in \hat{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}) \times \hat{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}) \times \mathbb{Z}_{\text{ad}}$  be the solution to the optimality system associated with the truncated optimal control problem defined in subsection 3.2 and  $(\bar{V}, \bar{P}, \bar{Z}) \in \mathbb{V}(\mathcal{T}_{\mathcal{Y}}) \times \mathbb{V}(\mathcal{T}_{\mathcal{Y}}) \times \mathbb{Z}_{\text{ad}}(\mathcal{T}_{\Omega})$  its numerical approximation defined in subsection 3.3. If (4.14) holds, then*

$$\begin{aligned} \|\nabla(\bar{v} - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_{\mathcal{Y}})} + \|\nabla(\bar{p} - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_{\mathcal{Y}})} + \|\bar{r} - \bar{Z}\|_{L^2(\Omega)} \\ \lesssim \mathcal{E}_V(\bar{V}, \bar{Z}; \mathcal{N}(\mathcal{T}_{\Omega})) + \mathcal{E}_P(\bar{P}, \bar{V}; \mathcal{N}(\mathcal{T}_{\Omega})) + \mathcal{E}_Z(\bar{Z}, \bar{P}; \mathcal{T}_{\Omega}), \end{aligned} \quad (4.15)$$

where the hidden constant is independent of the continuous and discrete optimal variables, and the size of the elements in the meshes  $\mathcal{T}_{\Omega}$  and  $\mathcal{T}_{\mathcal{Y}}$ .

*Proof.* The proof involves six steps.

Step 1. With the definition (4.12) of the local error indicator  $\mathcal{E}_Z$  in mind, we define the auxiliary control  $\tilde{r} = \Pi(-\frac{1}{\mu} \text{tr}_{\Omega} \bar{P})$  and notice that it verifies

$$(\text{tr}_{\Omega} \bar{P} + \mu \tilde{r}, r - \tilde{r})_{L^2(\Omega)} \geq 0 \quad \forall r \in \mathbb{Z}_{\text{ad}}. \quad (4.16)$$

Then, an application of the triangle inequality yields

$$\|\bar{r} - \bar{Z}\|_{L^2(\Omega)} \leq \|\bar{r} - \tilde{r}\|_{L^2(\Omega)} + \|\tilde{r} - \bar{Z}\|_{L^2(\Omega)} \quad (4.17)$$

We notice that the second term on the right hand side of the previous inequality corresponds to the definition of the global indicator (4.11). Thus, it suffices to bound the first term, i.e.,  $\|\bar{r} - \tilde{r}\|_{L^2(\Omega)}$ .

Step 2. Set  $r = \tilde{r}$  in (3.5) and  $r = \bar{r}$  in (4.16). Adding the obtained inequalities we arrive at

$$\mu \|\bar{r} - \tilde{r}\|_{L^2(\Omega)}^2 \leq (\text{tr}_{\Omega}(\bar{p} - \bar{P}), \tilde{r} - \bar{r})_{L^2(\Omega)}, \quad (4.18)$$

where  $\bar{p}$  and  $\bar{P}$  solve (3.6) and (3.15), respectively. To control the right hand side of this expression, we introduce the auxiliary adjoint state  $q$  that uniquely solves

$$q \in \hat{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}) : \quad a_{\mathcal{Y}}(\phi, q) = (\text{tr}_{\Omega} \bar{V} - u_{\text{d}}, \text{tr}_{\Omega} \phi)_{L^2(\Omega)} \quad \forall \phi \in \hat{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}). \quad (4.19)$$

By writing  $\bar{p} - \bar{P} = (\bar{p} - q) + (q - \bar{P})$ , the estimate (4.18) immediately yields

$$\mu \|\bar{r} - \tilde{r}\|_{L^2(\Omega)}^2 \leq (\text{tr}_{\Omega}(\bar{p} - q), \tilde{r} - \bar{r})_{L^2(\Omega)} + (\text{tr}_{\Omega}(q - \bar{P}), \tilde{r} - \bar{r})_{L^2(\Omega)}. \quad (4.20)$$

We conclude this step by noticing that, by construction, problem (3.15) corresponds to the Galerkin approximation of (4.19). Then, [22, Proposition 5.14] yields

$$\begin{aligned} |\text{II}| := |(\text{tr}_{\Omega}(q - \bar{P}), \tilde{r} - \bar{r})_{L^2(\Omega)}| &\lesssim \|\nabla(q - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_{\mathcal{Y}})} \|\tilde{r} - \bar{r}\|_{L^2(\Omega)} \\ &\lesssim \mathcal{E}_P(\bar{P}, \bar{V}; \mathcal{N}(\mathcal{T}_{\Omega})) \|\tilde{r} - \bar{r}\|_{L^2(\Omega)} \leq \frac{\mu}{4} \|\tilde{r} - \bar{r}\|_{L^2(\Omega)}^2 + C \mathcal{E}_P(\bar{P}, \bar{V}; \mathcal{N}(\mathcal{T}_{\Omega})), \end{aligned} \quad (4.21)$$

where in the first inequality we used (2.2);  $C$  denotes a positive constant.

Step 3. The goal of this step is to bound the term  $I := (\text{tr}_\Omega(\bar{p} - q), \tilde{r} - \bar{r})_{L^2(\Omega)}$ . To accomplish this task, we introduce another auxiliary adjoint state

$$w \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma) : \quad a_\gamma(\phi, w) = (\text{tr}_\Omega \tilde{v} - u_d, \text{tr}_\Omega \phi)_{L^2(\Omega)} \quad \forall \phi \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma), \quad (4.22)$$

where  $\tilde{v}$  is defined as the unique solution to

$$\tilde{v} \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma) : \quad a_\gamma(\tilde{v}, \phi) = (\tilde{r}, \text{tr}_\Omega \phi)_{L^2(\Omega)} \quad \forall \phi \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma), \quad (4.23)$$

and  $\tilde{r} = \Pi(-\frac{1}{\mu} \text{tr}_\Omega \bar{P})$ . We then write  $\bar{p} - q = (\bar{p} - w) + (w - q)$  and bound each contribution to the term  $I$  separately. To do this, we observe that  $\bar{v} - \tilde{v}$  solves the problem  $a_\gamma(\bar{v} - \tilde{v}, \phi) = (\bar{r} - \tilde{r}, \text{tr}_\Omega \phi)_{L^2(\Omega)}$  for all  $\phi \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma)$ . On the other hand, for all these test functions,  $\bar{p} - w$  solves  $a_\gamma(\phi, \bar{p} - w) = (\text{tr}_\Omega(\bar{v} - \tilde{v}), \text{tr}_\Omega \phi)_{L^2(\Omega)}$ . Combining these two problems, we arrive at

$$I_1 := (\text{tr}_\Omega(\bar{p} - w), \tilde{r} - \bar{r})_{L^2(\Omega)} = -a_\gamma(\bar{v} - \tilde{v}, \bar{p} - w) = -\|\text{tr}_\Omega(\bar{v} - \tilde{v})\|_{L^2(\Omega)}^2 \leq 0. \quad (4.24)$$

We now estimate the term  $I_2 := (\text{tr}_\Omega(w - q), \tilde{r} - \bar{r})_{L^2(\Omega)}$ , where  $w$  and  $q$  solve problems (4.22) and (4.19), respectively. We observe that the difference  $w - q$  solves  $a_\gamma(\phi, w - q) = (\text{tr}_\Omega(\tilde{v} - \bar{V}), \text{tr}_\Omega \phi)_{L^2(\Omega)}$  for all  $\phi \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma)$ . Thus, the trace estimate (2.2) and the stability of problem (4.19) yield

$$|I_2| \lesssim \|\nabla(w - q)\|_{L^2(y^\alpha, \mathcal{C}_\gamma)} \|\tilde{r} - \bar{r}\|_{L^2(\Omega)} \lesssim \|\text{tr}_\Omega(\tilde{v} - \bar{V})\|_{L^2(\Omega)} \|\tilde{r} - \bar{r}\|_{L^2(\Omega)}. \quad (4.25)$$

It suffices to bound the term  $\|\text{tr}_\Omega(\tilde{v} - \bar{V})\|_{L^2(\Omega)}$ . To accomplish this task, we invoke the triangle inequality and obtain the estimate  $\|\text{tr}_\Omega(\tilde{v} - \bar{V})\|_{L^2(\Omega)} \leq \|\text{tr}_\Omega(\tilde{v} - v^*)\|_{L^2(\Omega)} + \|\text{tr}_\Omega(v^* - \bar{V})\|_{L^2(\Omega)}$ , where  $v^*$  denotes the unique solution to the following problem:

$$v^* \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma) : \quad a_\gamma(v^*, \phi) = (\bar{Z}, \text{tr}_\Omega \phi)_{L^2(\Omega)} \quad \forall \phi \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\gamma). \quad (4.26)$$

Now, we invoke (2.2) and the stability of (4.26) to derive that  $\|\text{tr}_\Omega(\tilde{v} - v^*)\|_{L^2(\Omega)} \lesssim \|\tilde{r} - \bar{Z}\|_{L^2(\Omega)}$ . This, in view of the definition of  $\mathcal{E}_Z$ , given by (4.11)–(4.12), yields

$$\|\text{tr}_\Omega(\tilde{v} - v^*)\|_{L^2(\Omega)} \lesssim \mathcal{E}_Z(\bar{Z}, \bar{P}; \mathcal{T}_\Omega). \quad (4.27)$$

To control the remainder term, we observe that problem (3.13) corresponds to the Galerkin approximation of (4.26). Consequently, (2.2) and [22, Proposition 5.14] yield

$$\|\text{tr}_\Omega(v^* - \bar{V})\|_{L^2(\Omega)} \lesssim \|\nabla(v^* - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_\gamma)} \lesssim \mathcal{E}_V(\bar{V}, \bar{Z}; \mathcal{N}(\mathcal{T}_\Omega)). \quad (4.28)$$

In view of (4.25), the collection of the estimates (4.27) and (4.28) allows us to obtain

$$|I_2| \leq \frac{\mu}{4} \|\tilde{r} - \bar{r}\|_{L^2(\Omega)}^2 + C (\mathcal{E}_Z^2(\bar{Z}, \bar{P}; \mathcal{T}_\Omega) + \mathcal{E}_V^2(\bar{V}, \bar{Z}; \mathcal{N}(\mathcal{T}_\Omega))),$$

where  $C$  denotes a positive constant. Since (4.24) tells us that  $I_1 \leq 0$ , we obtain a similar estimate for the term  $I = I_1 + I_2$ . This estimate implies, on the basis of (4.20) and (4.21), the following bound

$$\|\bar{r} - \tilde{r}\|_{L^2(\Omega)}^2 \lesssim \mathcal{E}_V^2(\bar{V}, \bar{Z}; \mathcal{N}(\mathcal{T}_\Omega)) + \mathcal{E}_P^2(\bar{P}, \bar{V}; \mathcal{T}_\Omega) + \mathcal{E}_Z^2(\bar{Z}, \bar{P}; \mathcal{T}_\Omega),$$

which, invoking (4.17), provides an estimate for the error in control approximation:

$$\|\bar{r} - \bar{Z}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{T}_\gamma). \quad (4.29)$$

Step 4. The goal of this step is to bound the error  $\|\nabla(\bar{v} - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_\gamma)}$  in terms of the ideal error indicator (4.5). We employ similar arguments to the ones developed in step 2. We write  $\bar{v} - \bar{V} = (\bar{v} - v^*) + (v^* - \bar{V})$ , where  $v^*$  is defined in (4.26). The stability of problem (4.26) and the estimate (4.29) immediately provide the bound

$$\|\nabla(\bar{v} - v^*)\|_{L^2(y^\alpha, \mathcal{C}_\gamma)} \lesssim \|\bar{r} - \bar{Z}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{T}_\gamma).$$

which, combined with (4.28), allows us to derive

$$\|\nabla(\bar{v} - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_\gamma)} \lesssim \mathcal{E}_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{T}_\gamma). \quad (4.30)$$

Step 5. We bound the term  $\|\nabla(\bar{p} - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_\gamma)}$ . To accomplish this task, we invoke the triangle inequality and write

$$\|\nabla(\bar{p} - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_\gamma)} \leq \|\nabla(\bar{p} - q)\|_{L^2(y^\alpha, \mathcal{C}_\gamma)} + \|\nabla(q - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_\gamma)}.$$

where  $q$  is defined as the solution to problem (4.19). Applying the stability of problem (4.19), the trace estimate (2.2), and (4.30), we arrive at

$$\|\nabla(\bar{p} - q)\|_{L^2(y^\alpha, \mathcal{C}_\gamma)} \lesssim \|\text{tr}_\Omega(\bar{v} - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_\gamma)} \lesssim \mathcal{E}_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{T}_\gamma).$$

On the other hand, since  $\bar{P}$ , solution to (3.15), corresponds to the Galerkin approximation of  $q$ , solution to (4.19), we invoke [22, Proposition 5.14] to derive

$$\|\nabla(q - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_\gamma)} \lesssim \mathcal{E}_P(\bar{P}, \bar{Z}; \mathcal{K}(\mathcal{T}_\Omega)).$$

Collecting the derived estimates, we obtain that

$$\|\nabla(\bar{p} - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_\gamma)} \lesssim \mathcal{E}_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{T}_\gamma). \quad (4.31)$$

Step 6. Finally, the desired estimate (4.15) follows from a simple collection of the estimates (4.29), (4.30) and (4.31).  $\square$

We now derive a local lower bound that measures the quality of  $\mathcal{E}_{\text{ocp}}$  (efficiency). To achieve this, we define

$$C(d_s, \mu) = \max\{2d_s^{-1}, d_s^{-\frac{1}{2}}(\mu^{-1} + d_s^{-\frac{1}{2}}), 1 + d_s^{-\frac{1}{2}}\}. \quad (4.32)$$

**THEOREM 4.3 (local lower bound).** *Let  $(\bar{v}, \bar{p}, \bar{r}) \in \hat{H}_L^1(y^\alpha, \mathcal{C}_\gamma) \times \hat{H}_L^1(y^\alpha, \mathcal{C}_\gamma) \times \mathbf{Z}_{ad}$  be the solution to the optimality system associated with the truncated optimal control problem defined in subsection 3.2 and  $(\bar{V}, \bar{P}, \bar{Z}) \in \mathbb{V}(\mathcal{T}_\gamma) \times \mathbb{V}(\mathcal{T}_\gamma) \times \mathbb{Z}_{ad}(\mathcal{T}_\Omega)$  its numerical approximation defined in subsection 3.3. Then,*

$$\begin{aligned} & \mathcal{E}_V(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) + \mathcal{E}_P(\bar{P}, \bar{V}; \mathcal{C}_{z'}) + \mathcal{E}_Z(\bar{Z}, \bar{P}; \mathcal{S}_{z'}) \\ & \leq C(d_s, \mu) (\|\nabla(\bar{v} - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_{z'})} + \|\nabla(\bar{p} - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_{z'})} + \|\bar{r} - \bar{Z}\|_{L^2(\mathcal{S}_{z'})}), \end{aligned} \quad (4.33)$$

where  $C(d_s, \mu)$  depends only on  $d_s$  and the parameter  $\mu$  and is defined in (4.32).

*Proof.* We proceed in three steps.

Step 1. We begin by analyzing the efficiency properties of the indicator  $\mathcal{E}_V$  defined, locally, by (4.8). Let  $z' \in \mathcal{K}(\mathcal{T}_\Omega)$ . We invoke the fact that  $\zeta_{z'}$  solves the local problem (4.7) and conclude that

$$\mathcal{E}_V^2(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) = a_{z'}(\zeta_{z'}, \zeta_{z'}) = \langle \bar{r}, \text{tr}_\Omega \zeta_{z'} \rangle + \langle \bar{Z} - \bar{r}, \text{tr}_\Omega \zeta_{z'} \rangle - a_{z'}(\bar{V}, \zeta_{z'}).$$

Define  $e_V = \bar{v} - \bar{V}$ , where  $\bar{v}$  solves (3.4). Invoking (4.4) with  $C_{\text{tr}\Omega} \leq d_s^{-\frac{1}{2}}$  and a simple application of the Cauchy-Schwarz inequality, we arrive at

$$\begin{aligned} \mathcal{E}_V^2(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) &\leq d_s^{-1} \|\nabla e_V\|_{L^2(y^\alpha, \mathcal{C}_{z'})} \|\nabla \zeta_{z'}\|_{L^2(y^\alpha, \mathcal{C}_{z'})} + \|\bar{r} - \bar{Z}\|_{L^2(S_{z'})} \|\text{tr}_\Omega \zeta_{z'}\|_{L^2(S_{z'})} \\ &\leq \left( d_s^{-1} \|\nabla e_V\|_{L^2(y^\alpha, \mathcal{C}_{z'})} + d_s^{-\frac{1}{2}} \|\bar{r} - \bar{Z}\|_{L^2(S_{z'})} \right) \|\nabla \zeta_{z'}\|_{L^2(y^\alpha, \mathcal{C}_{z'})}. \end{aligned}$$

This, in view of definition (4.8), implies the efficiency of  $\mathcal{E}_V$ :

$$\mathcal{E}_V(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) \leq d_s^{-1} \|\nabla e_V\|_{L^2(y^\alpha, \mathcal{C}_{z'})} + d_s^{-\frac{1}{2}} \|\bar{r} - \bar{Z}\|_{L^2(S_{z'})}. \quad (4.34)$$

Step 2. In this step we elucidate the efficiency properties of the indicator  $\mathcal{E}_P$  defined in (4.10). Following the arguments elaborated in step 1, we write

$$\mathcal{E}_P^2(\bar{P}, \bar{V}; \mathcal{C}_{z'}) = \langle \text{tr}_\Omega(\bar{V} - \bar{v}), \chi_{z'} \rangle + a_{z'}(e_P, \chi_{z'}),$$

where  $\chi_{z'} \in \mathbb{W}(\mathcal{C}_{z'})$  solves (4.9). An application of (4.4) with  $C_{\text{tr}\Omega} \leq d_s^{-\frac{1}{2}}$  and the Cauchy-Schwarz inequality yield

$$\mathcal{E}_P(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) \leq d_s^{-1} \|\nabla e_V\|_{L^2(y^\alpha, \mathcal{C}_{z'})} + d_s^{-1} \|\nabla e_P\|_{L^2(y^\alpha, \mathcal{C}_{z'})}. \quad (4.35)$$

Step 3. The goal of this step is to analyze the efficiency properties of the indicator  $\mathcal{E}_Z$  defined by (4.11)–(4.12). A trivial application of the triangle inequality yields

$$\mathcal{E}_Z(\bar{Z}, \bar{P}; S_{z'}) \leq \|\bar{Z} - \Pi(-\frac{1}{\mu} \text{tr}_\Omega \bar{p})\|_{L^2(S_{z'})} + \|\Pi(-\frac{1}{\mu} \text{tr}_\Omega \bar{p}) - \Pi(-\frac{1}{\mu} \text{tr}_\Omega \bar{P})\|_{L^2(S_{z'})},$$

where  $\Pi$  denotes the nonlinear projector defined by (4.13). Now, in view of the local Lipschitz continuity of  $\Pi$ , the fact that  $\bar{r} = \Pi(-\frac{1}{\mu} \text{tr}_\Omega \bar{p})$  and the trace estimate (4.4) imply that

$$\mathcal{E}_Z(\bar{Z}, \bar{P}; S_{z'}) \leq \|\bar{r} - \bar{Z}\|_{L^2(S_{z'})} + \frac{d_s^{-\frac{1}{2}}}{\mu} \|\nabla e_P\|_{L^2(y^\alpha, \mathcal{C}_{z'})}. \quad (4.36)$$

Step 4. The desired estimate (4.33) follows from a collection of the estimates (4.34), (4.35), and (4.36). This concludes the proof.  $\square$

REMARK 4.4 (Local efficiency). Examining the proof of Theorem 4.3, we realize that the error indicators  $\mathcal{E}_V$ ,  $\mathcal{E}_P$  and  $\mathcal{E}_Z$  are locally efficient; see inequalities (4.34), (4.35) and (4.36), respectively. In addition, in all these inequalities the involved constants are known and depend only on the parameter  $s$ , through the constant  $d_s$ , and the parameter  $\mu$ . The key ingredients to derive the local efficiency property of the error estimator  $\mathcal{E}_Z$  are the local Lipschitz continuity of  $\Pi$  and the trace estimate (4.4). We comment that obtaining local a posteriori error bounds for the discretization of an optimal control problem is not always possible. We refer the reader to [42, Remark 3.3] for a thorough discussion on this matter.

**4.4. A computable a posteriori error estimator.** The a posteriori error estimator proposed and analyzed in subsection 4.3 has an obvious drawback: given a node  $z'$ , its construction requires the knowledge of the functions  $\zeta_{z'}$  and  $\chi_{z'}$  that solve exactly the infinite-dimensional problems (4.7) and (4.9), respectively. However, it provides intuition and sets the mathematical framework under which we will define a

computable and anisotropic a posteriori error estimator. To describe it, we define the following discrete local spaces.

DEFINITION 4.5 (discrete local spaces). *For  $z' \in \mathcal{N}(\mathcal{T}_\Omega)$ , we define*

$$\mathcal{W}(\mathcal{C}_{z'}) = \{W \in C^0(\overline{\mathcal{C}_{z'}}) : W|_T \in \mathcal{P}_2(K) \otimes \mathbb{P}_2(I) \ \forall T = K \times I \in \mathcal{C}_{z'}, \\ W|_{\partial \mathcal{C}_{z'} \setminus \Omega \times \{0\}} = 0\},$$

where, if  $K$  is a quadrilateral,  $\mathcal{P}_2(K)$  stands for  $\mathbb{Q}_2(K)$  — the space of polynomials of degree not larger than 2 in each variable. If  $K$  is a simplex,  $\mathcal{P}_2(K)$  corresponds to  $\mathbb{P}_2(K) \oplus \mathbb{B}(K)$  where  $\mathbb{P}_2(K)$  stands for the space of polynomials of total degree at most 2, and  $\mathbb{B}(K)$  is the space spanned by a local cubic bubble function.

With these discrete spaces at hand, we proceed to define the computable counterpart of the error indicator  $\mathcal{E}_{\text{ocp}}$  given by (4.5). This indicator is defined as follows:

$$E_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{T}_\Omega) = E_V(\bar{V}, \bar{Z}; \mathcal{N}(\mathcal{T}_\Omega)) + E_P(\bar{P}, \bar{V}; \mathcal{N}(\mathcal{T}_\Omega)) + E_Z(\bar{Z}, \bar{P}; \mathcal{T}_\Omega), \quad (4.37)$$

where  $\mathcal{T}_\Omega \in \mathbb{T}$  is the anisotropic mesh defined in subsection 3.3 and  $\bar{V}$ ,  $\bar{P}$  and  $\bar{Z}$  denote the optimal variables solving the fully discrete optimal control problem. To describe the first contribution in (4.37), we define  $\eta_{z'} \in \mathcal{W}(\mathcal{C}_{z'})$  as the solution to

$$a_{z'}(\eta_{z'}, W) = \langle \bar{Z}, \text{tr}_\Omega W \rangle - a_{z'}(\bar{V}, W) \quad \forall W \in \mathcal{W}(\mathcal{C}_{z'}). \quad (4.38)$$

We then define the local and computable error estimator, for the state equation, as

$$E_V(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) := \|\nabla \eta_{z'}\|_{L^2(y^\alpha, \mathcal{C}_{z'})}, \quad (4.39)$$

and the global error estimator  $E_V(\bar{V}, \bar{Z}; \mathcal{N}(\mathcal{T}_\Omega)) := \left( \sum_{z' \in \mathcal{N}(\mathcal{T}_\Omega)} E_V^2(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) \right)^{\frac{1}{2}}$ .

The second contribution in (4.37) is defined on the basis of the discrete object  $\theta_{z'} \in \mathcal{W}(\mathcal{C}_{z'})$  that solves the following local problem:

$$a_{z'}(\theta_{z'}, W) = \langle \text{tr}_\Omega \bar{V} - u_d, \text{tr}_\Omega W \rangle - a_{z'}(\bar{P}, W) \quad \forall W \in \mathcal{W}(\mathcal{C}_{z'}). \quad (4.40)$$

We thus define the local and computable error indicator

$$E_P(\bar{P}, \bar{V}; \mathcal{C}_{z'}) := \|\nabla \theta_{z'}\|_{L^2(y^\alpha, \mathcal{C}_{z'})} \quad (4.41)$$

and the global error indicator  $E_P(\bar{P}, \bar{V}; \mathcal{N}(\mathcal{T}_\Omega)) := \left( \sum_{z' \in \mathcal{N}(\mathcal{T}_\Omega)} E_P^2(\bar{P}, \bar{V}; \mathcal{C}_{z'}) \right)^{\frac{1}{2}}$ .

The third contribution in (4.37), i.e., the error indicator associated to the optimal control  $E_Z$ , is defined by (4.12)–(4.11).

We now explore the connection between the error estimator  $E_{\text{ocp}}$  and the error. We first obtain a lower bound that does not involve any oscillation term.

THEOREM 4.6 (local lower bound). *Let  $(\bar{v}, \bar{p}, \bar{r}) \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\Omega) \times \dot{H}_L^1(y^\alpha, \mathcal{C}_\Omega) \times \mathbb{Z}_{ad}$  be the solution to the optimality system associated with the truncated optimal control problem defined in subsection 3.2 and  $(\bar{V}, \bar{P}, \bar{Z}) \in \mathbb{V}(\mathcal{T}_\Omega) \times \mathbb{V}(\mathcal{T}_\Omega) \times \mathbb{Z}_{ad}(\mathcal{T}_\Omega)$  its numerical approximation defined in subsection 3.3. Then,*

$$E_V(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) + E_P(\bar{P}, \bar{V}; \mathcal{C}_{z'}) + E_Z(\bar{Z}, \bar{P}; \mathcal{S}_{z'}) \\ \leq C(d_s, \mu) \left( \|\nabla(\bar{v} - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_{z'})} + \|\nabla(\bar{p} - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_{z'})} + \|\bar{r} - \bar{Z}\|_{L^2(\mathcal{S}_{z'})} \right), \quad (4.42)$$

where  $C(d_s, \mu)$  depends only on  $d_s$  and the parameter  $\mu$  and is defined in (4.32).



*Proof.* The proof of the estimate (4.42) repeats the arguments developed in the proof of Theorem 4.3. We analyze the local efficiency of the indicator  $E_V$  defined in (4.39). To do this, we let  $z' \in \mathcal{N}(\mathcal{T}_\Omega)$ . Employing the fact that  $\eta_{z'}$  solves problem (4.38) and recalling that  $\bar{r}$  denotes the optimal control, we arrive at

$$E_V^2(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) = a_{z'}(\eta_{z'}, \eta_{z'}) = \langle \bar{r}, \eta_{z'} \rangle + \langle \bar{Z} - \bar{r}, \eta_{z'} \rangle - a_{z'}(\bar{V}, \eta_{z'}).$$

Invoking the trace estimate (4.4) with constant  $C_{\text{tr}\Omega} \leq d_s^{-\frac{1}{2}}$ , the fact that  $\bar{v}$  solves problem (3.4) and the Cauchy-Schwarz inequality, we obtain

$$E_V^2(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) \leq \left( d_s^{-1} \|\nabla(\bar{v} - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_{z'})} + d_s^{-\frac{1}{2}} \|\bar{r} - \bar{Z}\|_{L^2(S_{z'})} \right) \|\nabla \eta_{z'}\|_{L^2(y^\alpha, \mathcal{C}_{z'})},$$

which, in light of (4.39), immediately yields the desired result

$$E_V(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) \leq d_s^{-1} \|\nabla(\bar{v} - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_{z'})} + d_s^{-\frac{1}{2}} \|\bar{r} - \bar{Z}\|_{L^2(S_{z'})}.$$

The efficiency analysis for the terms  $E_P$  and  $E_Z$  follow similar arguments. For brevity we skip the proof.  $\square$

**REMARK 4.7** (strong efficiency). We remark that the lower bound (4.42) implies a strong concept of efficiency: it is free of any oscillation term and the involved constant  $C(d_s, \mu)$  is known and given by (4.32). The relative size of the local error indicator dictates mesh refinement regardless of fine structure of the data. The analysis is valid for the family of anisotropic meshes  $\mathcal{T}_\gamma$  and allows the nonuniformly coefficients involved in problems (3.4) and (3.6).

We now proceed to analyze the reliability properties of the anisotropic and computable error indicator  $E_{\text{ocp}}$  defined in (4.37). To achieve this, we introduce the so-called *data oscillation*. Given a function  $f \in L^2(\Omega)$  and  $z' \in \mathcal{N}(\mathcal{T}_\Omega)$ , we define the local oscillation of the function  $f$  as

$$\text{osc}(f; S_{z'}) := h_{z'}^s \|f - f_{z'}\|_{L^2(S_{z'})}, \quad (4.43)$$

where,  $h_{z'} = \min\{h_K : K \ni z'\}$  and  $f_{z'}|_K \in \mathbb{R}$  is the average of  $f$  over  $K$ , i.e.,

$$f_{z'}|_K := \fint_K f. \quad (4.44)$$

The global data oscillation is then defined as

$$\text{osc}(f; \mathcal{T}_\Omega) := \left( \sum_{z' \in \mathcal{N}(\mathcal{T}_\Omega)} \text{osc}(f; S_{z'})^2 \right)^{\frac{1}{2}}. \quad (4.45)$$

To present our results in a concise manner, we define  $D = (u_d, \text{tr}_\Omega \bar{V})$  and

$$\text{osc}(D; S_{z'}) := \text{osc}(u_d; S_{z'}) + \text{osc}(\text{tr}_\Omega \bar{V}; S_{z'}). \quad (4.46)$$

We define  $\text{osc}(D; \mathcal{T}_\Omega)$  accordingly. We also define the *total error indicator*

$$\mathcal{E}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{C}_{z'}) := (E_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{C}_{z'})^2 + \text{osc}(D; S_{z'})^2)^{\frac{1}{2}} \quad \forall z' \in \mathcal{N}(\mathcal{T}_\Omega). \quad (4.47)$$

This indicator will be used to mark elements for refinement in the AFEM proposed in section 5. The following remark is then necessary.

**REMARK 4.8** (marking). We comment that, in contrast to [19], the proposed AFEM will utilize the total error indicator, namely the sum of energy error and

oscillation, for marking. This could be avoided if  $E_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{C}_{z'}) \geq C \text{osc}(\mathbf{u}_d; S_{z'})$  for  $C > 0$ . While this property is trivial for the residual estimator with  $C = 1$ , it is in general false for other families of estimators such as the one we are proposing in this work. We refer to [19] for a thorough discussion on this matter.

Let  $\mathcal{H}_{\mathcal{T}_\Omega} = \{S_{z'} : z' \in \mathcal{N}(\mathcal{T}_\Omega)\}$  and, for any  $\mathcal{M} \subset \mathcal{H}_{\mathcal{T}_\Omega}$ , we set  $\mathcal{M}_{\mathcal{Y}} = \mathcal{M} \times (0, \mathcal{Y})$  and

$$\mathcal{E}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{M}_{\mathcal{Y}}) := \left( \sum_{S_{z'} \in \mathcal{M}} \mathcal{E}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{C}_{z'})^2 \right)^{1/2}, \quad (4.48)$$

where, we recall that  $\mathcal{C}_{z'} = S_{z'} \times (0, \mathcal{Y})$ . With these ingredients at hand, we present the following result.

**THEOREM 4.9** (global upper bound). *Let  $(\bar{v}, \bar{p}, \bar{r}) \in \hat{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}) \times \hat{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}) \times \mathbf{Z}_{ad}$  be the solution to the optimality system associated with the truncated optimal control problem defined in subsection 3.2 and  $(\bar{V}, \bar{P}, \bar{Z}) \in \mathbb{V}(\mathcal{T}_{\mathcal{Y}}) \times \mathbb{V}(\mathcal{T}_{\mathcal{Y}}) \times \mathbf{Z}_{ad}(\mathcal{T}_\Omega)$  its numerical approximation defined in subsection 3.3. If (4.14) and [23, Conjecture 5.28] hold, then*

$$\|\nabla(\bar{v} - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_{\mathcal{Y}})} + \|\nabla(\bar{p} - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_{\mathcal{Y}})} + \|\bar{r} - \bar{Z}\|_{L^2(\Omega)} \lesssim \mathcal{E}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{T}_{\mathcal{Y}}), \quad (4.49)$$

where the hidden constant is independent of the continuous and discrete optimal variables and the size of the elements in the meshes  $\mathcal{T}_\Omega$  and  $\mathcal{T}_{\mathcal{Y}}$ .

*Proof.* The proof of the estimate (4.49) follows closely the arguments developed in the proof of Theorem 4.2; the difference being the use of the computable error indicator  $E_{\text{ocp}}$  instead of the ideal estimator  $\mathcal{E}_{\text{ocp}}$ . We start by bounding the error in the control approximation. Defining  $\tilde{r} = \Pi(-\frac{1}{\mu} \text{tr}_\Omega \bar{P})$ , estimate (4.17) implies that

$$\|\bar{r} - \bar{Z}\|_{L^2(\Omega)} \leq \|\bar{r} - \tilde{r}\|_{L^2(\Omega)} + E_Z(\bar{Z}, \bar{P}; \mathcal{T}_\Omega). \quad (4.50)$$

To control the remainder term, we invoke (4.20) with  $q$  defined by (4.19) and write

$$\mu \|\bar{r} - \tilde{r}\|_{L^2(\Omega)}^2 \leq (\text{tr}_\Omega(\bar{p} - q), \bar{r} - \tilde{r})_{L^2(\Omega)} + (\text{tr}_\Omega(q - \bar{P}), \bar{r} - \tilde{r})_{L^2(\Omega)} = \text{I} + \text{II}. \quad (4.51)$$

To control the term II, we invoke the fact that  $\bar{P}$ , solution of problem (3.15), corresponds to the Galerkin approximation of  $q$ , solution of problem (4.19). This, in view of [22, Theorem 5.37], yields

$$|\text{II}| \leq \frac{\mu}{4} \|\bar{r} - \tilde{r}\|_{L^2(\Omega)}^2 + C (E_P^2(\bar{P}, \bar{V}; \mathcal{N}(\mathcal{T}_\Omega)) + \text{osc}^2(D; \mathcal{T}_\Omega)), \quad (4.52)$$

where  $C$  denotes a positive constant and  $\text{osc}$  is defined by (4.43) and (4.45).

To control the term I, we write  $\text{I} = \text{I}_1 + \text{I}_2 := (\text{tr}_\Omega(\bar{p} - w), \bar{r} - \tilde{r})_{L^2(\Omega)} + (\text{tr}_\Omega(w - q), \bar{r} - \tilde{r})_{L^2(\Omega)}$ , where  $w$  is defined as the solution to (4.22). Step 3 in the proof of Theorem 4.2 implies that  $\text{I}_1 \leq 0$ . To control the term  $\text{I}_2$ , we invoke (4.25) and write

$$|\text{I}_2| \lesssim \|\nabla(w - q)\|_{L^2(y^\alpha, \mathcal{C}_{\mathcal{Y}})} \|\bar{r} - \tilde{r}\|_{L^2(\Omega)} \lesssim \|\text{tr}_\Omega(\tilde{v} - \bar{V})\|_{L^2(\Omega)} \|\bar{r} - \tilde{r}\|_{L^2(\Omega)}. \quad (4.53)$$

We now write  $\tilde{v} - \bar{V} = (\tilde{v} - v^*) - (v^* - \bar{V})$ , where  $v^*$  solves (4.26), and estimate each contribution separately. First, stability of (4.26) yields

$$\|\text{tr}_\Omega(\tilde{v} - v^*)\|_{L^2(\Omega)} \lesssim E_Z(\bar{Z}, \bar{P}; \mathcal{T}_\Omega). \quad (4.54)$$

Second, since  $\bar{V}$  corresponds to the Galerkin approximation of  $v^*$ , [22, Theorem 5.37] implies the estimate

$$\|\operatorname{tr}_\Omega(v^* - \bar{V})\|_{L^2(\Omega)} \lesssim \|\nabla(v^* - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_y)} \lesssim E_V(\bar{V}, \bar{Z}; \mathcal{N}(\mathcal{T}_\Omega)). \quad (4.55)$$

This, in view of (4.53) and (4.54), implies that

$$|I_2| \leq \frac{\mu}{4} \|\bar{r} - \bar{r}\|_{L^2(\Omega)}^2 + C(E_V^2(\bar{V}, \bar{Z}; \mathcal{N}(\mathcal{T}_\Omega)) + E_Z^2(\bar{Z}, \bar{P}; \mathcal{T}_\Omega)),$$

where  $C$  denotes a positive constant. Since  $I_1 \leq 0$ , a similar estimate holds for  $I = I_1 + I_2$ . This estimate, in conjunction with the previous bound, and the estimates (4.50), (4.52) and (4.51) imply that

$$\|\bar{r} - \bar{r}\|_{L^2(\Omega)}^2 \lesssim E_{\text{ocp}}^2(\bar{V}, \bar{P}, \bar{Z}; \mathcal{T}_\Omega) + \operatorname{osc}^2(D; \mathcal{T}_\Omega).$$

The estimates for the terms  $\|\nabla(\bar{v} - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_y)}$  and  $\|\nabla(\bar{p} - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_y)}$  follow similar arguments to the ones elaborated in the steps 4 and 5 of the proof of Theorem 4.2. For brevity we skip the details.  $\square$

**REMARK 4.10** (Conjecture 5.28 in [23]). Examining the proof of Theorem 4.9, we realize that the key steps where [22, Theorem 5.37] is invoked are (4.51) and (4.55). The results of [22, Theorem 5.37] are valid under the assumption of the existence of an operator  $\mathcal{M}_{z'}$  that verify the conditions stipulated in [22, Conjecture 5.28]. The construction of the operator  $\mathcal{M}_{z'}$  is an open problem. The numerical experiments of [23, section 6] provide consistent computational evidence of the existence of  $\mathcal{M}_{z'}$  with the requisite properties.

**5. Numerical Experiments.** In this section we conduct a numerical example that illustrates the performance of the proposed error estimator. To accomplish this task, we formulate an adaptive finite element method (AFEM) based on the following iterative loop:

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE} \quad (5.1)$$

**5.1. Design of AFEM.** We proceed to describe the four modules in (5.1)

- **SOLVE:** Given  $\mathcal{T}_y$ , we compute  $(\bar{Y}, \bar{P}, \bar{Z}) \in \mathbb{V}(\mathcal{T}_y) \times \mathbb{V}(\mathcal{T}_y) \times \mathbb{Z}_{ad}(\mathcal{T}_\Omega)$ , the solution to the fully discrete optimal control problem defined in subsection 3.3:

$$(\bar{Y}, \bar{P}, \bar{U}) = \text{SOLVE}(\mathcal{T}_y).$$

To solve the minimization problem, we have used the projected BFGS method with Armijo line search; see [41]. The optimization algorithm is terminated when the  $\ell^2$ -norm of the projected gradient is less or equal to  $10^{-5}$ .

- **ESTIMATE:** Once a discrete solution is obtained, we compute, for each  $z' \in \mathcal{N}(\mathcal{T}_\Omega)$ , the local error indicator (4.37), which reads

$$E_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{C}_{z'}) = E_V(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) + E_P(\bar{P}, \bar{V}; \mathcal{C}_{z'}) + E_Z(\bar{Z}, \bar{P}; \mathcal{S}_{z'}),$$

where the indicators  $E_V$ ,  $E_P$ ,  $E_Z$  are defined by (4.39), (4.41) and (4.12), respectively. We then compute the oscillation term (4.46) and construct the total error indicator (4.47):

$$\{\mathcal{E}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{S}_{z'})\}_{\mathcal{S}_{z'} \in \mathcal{K}_{\mathcal{T}_\Omega}} = \text{ESTIMATE}(\bar{V}, \bar{P}, \bar{Z}; \mathcal{T}_y),$$

where  $\mathcal{K}_{\mathcal{T}_\Omega} = \{\mathcal{S}_{z'} : z' \in \mathcal{N}(\mathcal{T}_\Omega)\}$ . For notational convenience, and in view of the fact that  $\mathcal{C}_{z'} = \mathcal{S}_{z'} \times (0, \gamma)$  we replaced  $\mathcal{C}_{z'}$  by  $\mathcal{S}_{z'}$  in the previous formula.

- **MARK:** Using the so-called Dörfler marking strategy [29] (bulk chasing strategy) with parameter  $\theta$  with  $\theta \in (0, 1]$ , we select a set

$$\mathcal{M} = \text{MARK} \left( \left\{ \mathcal{E}(\bar{V}, \bar{P}, \bar{Z}; S_{z'}) \right\}_{S_{z'} \in \mathcal{K}_{\mathcal{T}_\Omega}}, (\bar{V}, \bar{P}, \bar{Z}) \right) \subset \mathcal{K}_{\mathcal{T}_\Omega}$$

of minimal cardinality that satisfies

$$\mathcal{E}((\bar{V}, \bar{P}, \bar{Z}), \mathcal{M}) \geq \theta \mathcal{E}((\bar{V}, \bar{P}, \bar{Z}), \mathcal{K}_{\mathcal{T}_\Omega}).$$

- **REFINEMENT:** We generate a new mesh  $\mathcal{T}'_\Omega$  by bisecting all the elements  $K \in \mathcal{T}_\Omega$  contained in  $\mathcal{M}$  based on the newest-vertex bisection method [55, 56]. We choose the truncation parameter as  $\gamma = 1 + \frac{1}{3} \log(\#\mathcal{T}'_\Omega)$  to balance the approximation and truncation errors [53, Remark 5.5]. The mesh  $\mathcal{T}'_\gamma$  is constructed by the rule (3.10), with a number of degrees of freedom  $M$  sufficiently large so that (4.14) holds. This is attained by first creating a partition  $\mathcal{T}'_\gamma$  with  $M \approx (\#\mathcal{T}'_\Omega)^{1/n}$  and checking (4.14). If this condition is violated, we increase the number of points until we get the desired result. The new mesh

$$\mathcal{T}'_\gamma = \text{REFINE}(\mathcal{M}),$$

is obtained as the tensor product of  $\mathcal{T}'_\Omega$  and  $\mathcal{T}'_\gamma$ .

**5.2. Implementation.** The AFEM (5.1) is implemented within the MATLAB software library iFEM [21]. All matrices have been assembled exactly. The right hand sides are computed by a quadrature formula which is exact for polynomials of degree 4. All linear systems were solved using the multigrid method with line smoother introduced and analyzed in [23].

To compute the solution  $\eta_{z'}$  to the discrete local problem (4.38) we proceed as follow: we loop around each node  $z' \in \mathcal{N}(\mathcal{T}_\Omega)$ , collect data about the cylindrical star  $\mathcal{C}_{z'}$  and assemble the small linear system (4.38). This linear system is solved by the built-in *direct solver* of MATLAB. To compute the solution  $\theta_{z'}$  to the discrete local problem (4.40), we proceed similarly. All integrals involving only the weight and discrete functions are computed exactly, whereas those also involving data functions are computed element-wise by a quadrature formula which is exact for polynomials of degree 7.

For convenience, in the MARK step we change the estimator from star-wise to element-wise. To accomplish this task, we first scale the nodal-wise estimator as  $E_{\text{ocp}}^2(\bar{V}, \bar{P}, \bar{Z}; \mathcal{C}_{z'})/(\#S_{z'})$  and then, for each element  $K \in \mathcal{T}_\Omega$ , we compute

$$E_{\text{ocp}}^2(\bar{V}, \bar{P}, \bar{Z}; K_\gamma) := \sum_{z' \in K} E_{\text{ocp}}^2(\bar{V}, \bar{P}, \bar{Z}; \mathcal{C}_{z'}),$$

where  $K_\gamma = K \times (0, \gamma)$ . The scaling is introduced so that

$$\sum_{K \in \mathcal{T}_\Omega} E_{\text{ocp}}^2(\bar{V}, \bar{P}, \bar{Z}; K_\gamma) = \sum_{z' \in \mathcal{N}(\mathcal{T}_\Omega)} E_{\text{ocp}}^2(\bar{V}, \bar{P}, \bar{Z}; \mathcal{C}_{z'}).$$

The cell-wise data oscillation is now defined as

$$\text{osc}(f; K)^2 := h_K^{2s} \|f - \bar{f}_K\|_{L^2(K)}^2,$$

where  $\bar{f}_K$  denotes the average of  $f$  over the element  $K$ . This quantity is computed using a quadrature formula which is exact for polynomials of degree 7.

**5.3. L-shaped domain with incompatible data.** For our numerical example we consider the worst possible scenario:

- (D1)  $a = 0.1, b = 0.3$ . This implies that the optimal control  $\bar{z} \notin \mathbb{H}^{1-s}(\Omega)$  when  $s < \frac{1}{2}$ . We will refer to the optimal control  $\bar{z}$  as *incompatible datum* for (3.4).
- (D2)  $u_d = 1$ . This element does not belong to  $\mathbb{H}^{1-s}(\Omega)$  when  $s < \frac{1}{2}$ . Therefore, for  $s < \frac{1}{2}$ ,  $u_d$  is an *incompatible datum* for problem (3.6).
- (D3)  $\Omega = (-1, 1)^2 \setminus (0, 1) \times (-1, 0)$ , i.e., the well known L-shaped domain; see Figure 5.2.

In view of (D1) and (D2), we conclude that the right hand sides of the state and adjoint equations, problems (3.4) and (3.6), respectively, are incompatible for  $s < \frac{1}{2}$ . As discussed in [53, section 6.3], at the level of the state equation, this results in lower rates of convergence when quasi-uniform refinement of  $\Omega$  is employed. In addition, we consider a situation where the domain  $\Omega$  is nonconvex. As a result, the hypothesis of Theorem 3.17 does not hold and then it cannot be applied.

We set  $\mu = 1$ , and we comment that we do not explicitly enforce the mesh restriction (4.14), which shows that this is nothing but an artifact in our theory.

As Figure 5.1 illustrates, using our proposed AFEM driven by the error indicator (4.47), we can recover the optimal rates of convergence (3.17)–(3.18) for all values of  $s$  considered:  $s = 0.2, 0.4, 0.6$ , and  $s = 0.8$ . We remark, again, that we are operating under the conditions (D1)–(D3) and then Theorem 3.17 cannot be applied. Since, for  $s < \frac{1}{2}$ , the data is incompatible (D1)–(D2), the optimal and adjoint states exhibit boundary layers. To capture them, our AFEM refines near the boundary; see Figure 5.2 (middle). In contrast, when  $s > \frac{1}{2}$  such incompatibilities does not occur and then our AFEM focuses to resolve the reentrant corner; see Figure 5.2 (right). The left panel in Figure 5.2 shows the initial mesh. We comment that the middle and the right panels are obtained with 17 AFEM cycles.

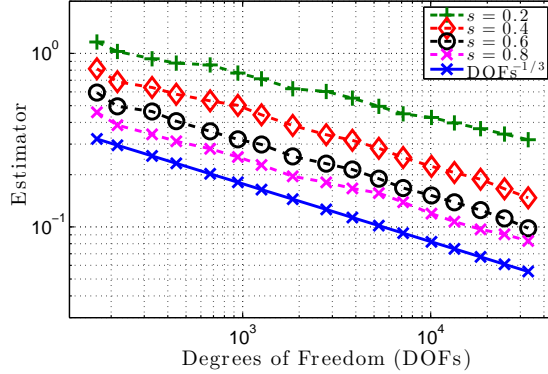


FIG. 5.1. Computational rate of convergence for our anisotropic AFEM with incompatible right hand sides for both the state equation and the adjoint equation over an L-shaped domain (non-convex domain). We consider  $n = 2$ . Since the exact solution is not known for this problem, we present the total error estimator with respect to the number of degrees of freedom. In all cases we recover the optimal rate of convergence  $(\#\mathcal{T}_\gamma)^{-1/3}$ .

## REFERENCES

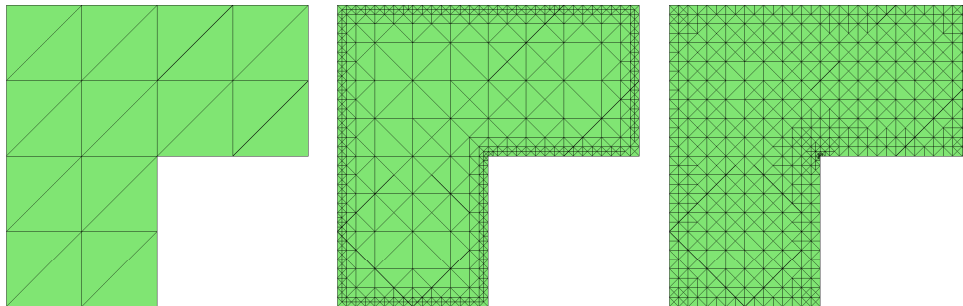


FIG. 5.2. The left panel shows the initial grid. The middle and right panels show adaptive grids, obtained after 17 refinements, for  $s = 0.2$  and  $s = 0.8$ , respectively. We consider an L-shaped domain with incompatible right hand side for the state and adjoint equations. As expected when  $s < \frac{1}{2}$  the incompatible data ( $\bar{z}, u_d \notin \mathbb{H}^{1-s}(\Omega)$ ) results in boundary layers for both the state and the adjoint state. In order to capture them, our AFEM refines near the boundary. In contrast, when  $s > \frac{1}{2}$  the refinement is more pronounced near the reentrant corner; the data  $\bar{z}$  and  $u_d$  are compatible in this case.

- [1] S. Abe and S. Thurner. Anomalous diffusion in view of Einstein's 1905 theory of Brownian motion. *Physica A: Statistical Mechanics and its Applications*, 356(24):403 – 407, 2005.
- [2] M. Abramowitz and I.A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. 1964.
- [3] M. Ainsworth and J.T. Oden. *A posteriori error estimation in finite element analysis*. Pure and Applied Mathematics (New York). Wiley-Interscience, New York, 2000.
- [4] H. Antil and E. Otárola. A FEM for an optimal control problem of fractional powers of elliptic operators. *SIAM J. Control Optim.*, 53(6):3432–3456, 2015.
- [5] N. Arada, E. Casas, and F. Tröltzsch. Error estimates for the numerical approximation of a semilinear elliptic control problem. *Comput. Optim. Appl.*, 23(2):201–229, 2002.
- [6] T.M. Atanackovic, S. Pilipovic, B. Stankovic, and D. Zorica. *Fractional Calculus with Applications in Mechanics: Vibrations and Diffusion Processes*. John Wiley & Sons, 2014.
- [7] I. Babuška and A. Miller. A feedback finite element method with a posteriori error estimation. I. The finite element method and some basic properties of the a posteriori error estimator. *Comput. Methods Appl. Mech. Engrg.*, 61(1):1–40, 1987.
- [8] I. Babuška and W.C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15(4):736–754, 1978.
- [9] I. Babuška and T. Strouboulis. *The Finite Element Method and Its Reliability*. Numerical mathematics and scientific computation. Clarendon Press, 2001.
- [10] E. Barkai, R. Metzler, and J. Klafter. From continuous time random walks to the fractional Fokker-Planck equation. *Phys. Rev. E* (3), 61(1):132–138, 2000.
- [11] R. Becker, H. Kapp, and R. Rannacher. Adaptive finite element methods for optimal control of partial differential equations: basic concept. *SIAM J. Control Optim.*, 39(1):113–132 (electronic), 2000.
- [12] J.-P. Bouchaud and A. Georges. Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications. *Phys. Rep.*, 195(4-5):127–293, 1990.
- [13] C. Bucur and E. Valdinoci. Nonlocal diffusion and applications. arXiv:1504.08292, 2015.
- [14] A. Bueno-Orovio, D. Kay, V. Grau, B. Rodriguez, and K. Burrage. Fractional diffusion models of cardiac electrical propagation: role of structural heterogeneity in dispersion of repolarization. *J. R. Soc. Interface*, 11(97), 2014.
- [15] L. Caffarelli and L. Silvestre. An extension problem related to the fractional Laplacian. *Comm. Part. Diff. Eqs.*, 32(7-9):1245–1260, 2007.
- [16] A. Capella, J. Dávila, L. Dupaigne, and Y. Sire. Regularity of radial extremal solutions for some non-local semilinear equations. *Comm. Part. Diff. Eqs.*, 36(8):1353–1384, 2011.
- [17] C. Carstensen and S.A. Funken. Fully reliable localized error control in the FEM. *SIAM J. Sci. Comput.*, 21(4):1465–1484 (electronic), 1999/00.
- [18] E. Casas, M. Mateos, and F. Tröltzsch. Error estimates for the numerical approximation of boundary semilinear elliptic control problems. *Comput. Optim. Appl.*, 31(2):193–219, 2005.

- [19] J.M. Cascón, C. Kreuzer, R.H. Nochetto, and K.G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.
- [20] J.M. Cascón and R.H. Nochetto. Quasioptimal cardinality of AFEM driven by nonresidual estimators. *IMA J. Numer. Anal.*, 32(1):1–29, 2012.
- [21] L. Chen. *iFEM: An integrated finite element methods package in matlab*. Technical report, University of California at Irvine, 2009.
- [22] L. Chen, R. H. Nochetto, E. Otárola, and A. J. Salgado. A PDE approach to fractional diffusion: a posteriori error analysis. *J. Comput. Phys.*, 293:339–358, 2015.
- [23] L. Chen, R.H. Nochetto, E. Otárola, and A.J. Salgado. Multilevel methods for nonuniformly elliptic operators and fractional diffusion. *Math. Comp.*, 2016. (to appear).
- [24] W. Chen. A speculative study of 2/3-order fractional laplacian modeling of turbulence: Some thoughts and conjectures. *Chaos*, 16(2):1–11, 2006.
- [25] P.G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. SIAM, Philadelphia, PA, 2002.
- [26] L. Debnath. Fractional integral and fractional differential equations in fluid mechanics. *Fract. Calc. Appl. Anal.*, 6(2):119–155, 2003.
- [27] L. Debnath. Recent applications of fractional calculus to science and engineering. *Int. J. Math. Math. Sci.*, (54):3413–3442, 2003.
- [28] D. del Castillo-Negrete, B. A. Carreras, and V. E. Lynch. Fractional diffusion in plasma turbulence. *Physics of Plasmas*, 11(8):3854–3864, 2004.
- [29] W. Dörfler. A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.
- [30] J. Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.
- [31] R.G. Durán and A.L. Lombardi. Error estimates on anisotropic  $Q_1$  elements for functions in weighted Sobolev spaces. *Math. Comp.*, 74(252):1679–1706 (electronic), 2005.
- [32] L. Formaggia and S. Perotto. New anisotropic a priori error estimates. *Numer. Math.*, 89(4):641–667, 2001.
- [33] L. Formaggia and S. Perotto. Anisotropic error estimates for elliptic problems. *Numer. Math.*, 94(1):67–92, 2003.
- [34] P. Gatto and J.S. Hesthaven. Numerical approximation of the fractional laplacian via hp-finite elements, with an application to image denoising. *J. Sci. Comp.*, 65(1):249–270, 2015.
- [35] V. Gol’dshstein and A. Ukhlov. Weighted Sobolev spaces and embedding theorems. *Trans. Amer. Math. Soc.*, 361(7):3829–3850, 2009.
- [36] R. Gorenflo, F. Mainardi, D. Moretti, and P. Paradisi. Time fractional diffusion: a discrete random walk approach. *Nonlinear Dynam.*, 29(1-4):129–143, 2002. Fractional order calculus and its applications.
- [37] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.
- [38] M. Hintermüller, R. H. W. Hoppe, Y. Iliash, and M. Kieweg. An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. *ESAIM: Control Optim. Calc. of Var.*, 14:540–560, 7 2008.
- [39] M. Hinze. A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. Appl.*, 30(1):45–61, 2005.
- [40] R. Ishizuka, S.-H. Chong, and F. Hirata. An integral equation theory for inhomogeneous molecular fluids: The reference interaction site model approach. *J. Chem. Phys.*, 128(3), 2008.
- [41] C. T. Kelley. *Iterative methods for optimization*, volume 18 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.
- [42] K. Kohls, A. Rösch, and K.G. Siebert. A posteriori error analysis of optimal control problems with control constraints. *SIAM J. Control Optim.*, 52(3):1832–1861, 2014.
- [43] G. Kunert. An a posteriori residual error estimator for the finite element method on anisotropic tetrahedral meshes. *Numer. Math.*, 86(3):471–490, 2000.
- [44] G. Kunert and R. Verfürth. Edge residuals dominate a posteriori error estimates for linear finite element methods on anisotropic triangular and tetrahedral meshes. *Numer. Math.*, 86(2):283–303, 2000.
- [45] N.S. Landkof. *Foundations of modern potential theory*. Springer-Verlag, New York, 1972. Translated from the Russian by A. P. Doohovskoy, Die Grundlehren der mathematischen Wissenschaften, Band 180.
- [46] S. Z. Levendorskiĭ. Pricing of the American put under Lévy processes. *Int. J. Theor. Appl. Finance*, 7(3):303–335, 2004.
- [47] S. Micheletti and S. Perotto. The effect of anisotropic mesh adaptation on PDE-constrained

- optimal control problems. *SIAM J. Control Optim.*, 49(4):1793–1828, 2011.
- [48] P. Morin, R. H. Nochetto, and K. G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488 (electronic), 2000.
  - [49] P. Morin, R.H. Nochetto, and K.G. Siebert. Local problems on stars: a posteriori error estimators, convergence, and performance. *Math. Comp.*, 72(243):1067–1097 (electronic), 2003.
  - [50] B. Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.*, 165:207–226, 1972.
  - [51] R.R. Nigmatullin. The realization of the generalized transfer equation in a medium with fractal geometry. *Physica Status Solidi (b)*, 133(1):425–430, 1986.
  - [52] R. H. Nochetto, E. Otárola, and A. J. Salgado. Piecewise polynomial interpolation in Muckenhoupt weighted Sobolev spaces and applications. *Numer. Math.*, 132(1):85–130, 2016.
  - [53] R.H. Nochetto, E. Otárola, and A. J. Salgado. A PDE approach to fractional diffusion in general domains: A priori error analysis. *Found. Comput. Math.*, 15(3):733–791, 2015.
  - [54] R.H. Nochetto, E. Otárola, and A.J. Salgado. A PDE approach to space-time fractional parabolic problems. *SIAM J. Numer. Anal.*, 54(2):848–873, 2016.
  - [55] R.H. Nochetto, K.G. Siebert, and A. Veeser. Theory of adaptive finite element methods: an introduction. In *Multiscale, nonlinear and adaptive approximation*. Springer, 2009.
  - [56] R.H. Nochetto and A. Vesser. Primer of adaptive finite element methods. In *Multiscale and Adaptivity: Modeling, Numerics and Applications, CIME Lectures*. Springer, 2011.
  - [57] M. Picasso. Anisotropic a posteriori error estimate for an optimal control problem governed by the heat equation. *Numer. Methods Partial Differential Equations*, 22(6):1314–1336, 2006.
  - [58] A. Weiser R. E. Bank. Some a posteriori error estimators for elliptic partial differential equations. *Math. Comp.*, 44(170):283–301, 1985.
  - [59] A.I. Saichev and G.M. Zaslavsky. Fractional kinetic equations: solutions and applications. *Chaos*, 7(4):753–764, 1997.
  - [60] K. G. Siebert. An a posteriori error estimator for anisotropic refinement. *Numer. Math.*, 73(3):373–398, 1996.
  - [61] L. Silvestre. Regularity of the obstacle problem for a fractional power of the Laplace operator. *Comm. Pure Appl. Math.*, 60(1):67–112, 2007.
  - [62] E.M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.
  - [63] P.R. Stinga and J.L. Torrea. Extension problem and Harnack’s inequality for some fractional operators. *Comm. Part. Diff. Eqs.*, 35(11):2092–2122, 2010.
  - [64] B.O. Turesson. *Nonlinear potential theory and weighted Sobolev spaces*, volume 1736 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000.
  - [65] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. John Wiley, 1996.